

Albrecht Schmidt, Paul Holleis
Matthias Kranz, Andreas Butz (Hrsg.)

Innovative Benutzungsschnittstellen und Neue Formen der Interaktion

Hauptseminar Medieninformatik WS 2004/05

Technical Report
LMU-MI-2005-1, Mar. 2005
ISSN 1862-5207



University of Munich
Department of Computer Science
Media Informatics Group

Albrecht Schmidt, Paul Holleis, Matthias Kranz, Andreas Butz (Herausgeber)

Innovative Benutzungsschnittstellen und Neue Formen der Interaktion

Ein Überblick über Forschungsarbeiten im Bereich der
Mensch-Maschine-Interaktion

Vorwort

In den letzten 10 Jahren haben sich massive Veränderungen im Bereich der Benutzungsschnittstellen vollzogen. Mit diesem Bericht über *Innovative Benutzungsschnittstellen und Neue Formen der Interaktion* wollen wir kompakt aktuelle Entwicklungen und Forschungstrends im Bereich Mensch-Maschine-Interaktion einem interessierten Fachpublikum zugänglich machen. Dazu analysieren Studenten und Mitarbeiter des Lehrstuhl Medieninformatik an der Ludwig-Maximilians-Universität München Beiträge und Veröffentlichungen aus zahlreichen Workshops, Konferenzen und wissenschaftlichen Zeitschriften. Diese Reihe an Texten wurde im Zusammenhang mit einem Hauptseminar im Wintersemester 2004/05 von Studenten erstellt.

Jedes der einzelnen Kapitel greift ein Thema auf und stellt wesentliche Forschungstrends in diesem Bereich in kurzen Artikeln in deutscher Sprache vor. Im Rahmen des Seminars wurde von den Studenten zu den jeweiligen Themen Vorträge gehalten, die die wesentlichen Inhalte noch einmal zusammenfassten. Die Folien zu diesen Vorträgen und weitere Informationen über die Veranstaltung können auf der Webseite <http://www.hcilab.org/events/ui-update1.0/> eingesehen werden.

Dieser Bericht richtet sich in erster Linie an Informatiker, Medieninformatiker, User Interface Designer und Studenten der Informatik und Medieninformatik. Betrachtet man den Trend auch Alltagsgegenstände (Kleidung, Kaffeemaschine, usw.) mit Technologie auszustatten und sie somit in Benutzungsschnittstellen zu verwandeln, erhält das Thema eine größere Tragweite. Im weiteren Sinne richtet sich der Bericht an alle die sich für neue Formen der Interaktion zwischen Mensch und Maschine interessieren.

Veranstalter und Herausgeber sind die beiden von der DFG geförderten Nachwuchsforschungsgruppen „Embedded Interaction“ (<http://www.hcilab.org>) und „Fluidum“ (<http://www.fluidum.org>) an der Ludwig-Maximilians-Universität München.

München, März 2005

Die Herausgeber

Albrecht Schmidt

Paul Holleis

Matthias Kranz

Andreas Butz

Inhaltsverzeichnis

<i>Ronald Ecker</i> Benutzerschnittstellen für erweiterte Realitäten	1
<i>Julius Bahr</i> Interaktive Oberflächen und Table Top Benutzerschnittstellen	15
<i>Matthias Schicker</i> Projizierte Benutzerschnittstellen und steuerbare Projektoren	35
<i>Florian Block</i> Umgebungsmedien und Stille Technologie.....	55
<i>Dagmar Kern</i> Gefühlsorientierte Benutzerschnittstellen.....	71
<i>Monika Wnuk</i> Benutzerschnittstellen zum Anziehen.....	87
<i>Martin Schrittenloher</i> Neue Formen der entfernten Kommunikation	105
<i>Raphael Wimmer</i> Sprachbasierte Benutzerschnittstellen	117
<i>Marc Breisinger</i> 3D Benutzungsschnittstellen, 3D Scanning und Rekonstruktion	137
<i>Benjamin Schlerf</i> Berührungsgesteuerte Benutzerschnittstellen.....	157
<i>Radostina Ruseva</i> Neue Benutzerschnittstellen für Spiele.....	173
<i>Stefan Freund</i> Berührbare Benutzerschnittstellen und Daten zum Anfassen	189
<i>Friederike Otto</i> Greifbare Programmierung	207

Interaktive Oberflächen und “Table Top” Benutzerschnittstellen

Julius Bahr

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
dschulius@mac.com

Zusammenfassung Das vorliegende Paper beschäftigt sich mit den physikalischen Grundlagen, der Zusammensetzung und dem Aufbau interaktiver Oberflächen und “Table Top” Benutzerschnittstellen. In der Einführung wird die Nutzung interaktiver Oberflächen und “Table Top” Benutzerschnittstellen motiviert. Im weiteren werden Anwendungsbereiche anhand von Beispielen vorgestellt. Besonderes Augenmerk wird der Interaktion mit Interaktiven Oberflächen und “Table Top” Benutzerschnittstellen zukommen. Die Arbeit schliesst mit einem Ausblick auf mögliche, zukünftige Anwendungen und einer Kritik des heutigen Forschungsstand.

1 Einführung

Interaktive Oberflächen & “Table-Top” Benutzerschnittstellen sind ein Forschungsgebiet im Bereich der Erweiterten Realität (Augmented Reality). Das Ziel ist es möglicherweise ungenutzte Flächen interaktiv zu machen und dadurch mit der Umwelt in Kontakt zu treten und zu interagieren. Ein Beispiel hierfür sind Werbeflächen die interaktiv sind. In einem Autohaus könnte die Schaufensterscheibe interaktiv sein. Auf der Scheibe wird dann eine Anwendung z.B. zur Fahrzeugkonfiguration laufen. Mittels Radar wäre es sogar möglich festzustellen, wenn jemand vor der Scheibe steht und den potenziellen Kunden direkt anzusprechen. Siehe auch [1] S.4 und Abbildung 1. Ein typisches Charakteristikum interaktiver Oberflächen ist die Benutzung der Hand als Eingabemedium. Es gibt keine Maus, keine Tastatur.

Dieses Paper beschränkt sich auf feststehende interaktive Oberflächen. Im weiteren wird für Interaktiver Oberflächen und “Table Top” Benutzerschnittstellen die Abkürzung IO verwendet.

2 Merkmale und genereller Aufbau interaktiver Oberflächen

Eine interaktive Oberfläche setzt sich aus 3 Elementen zusammen: Einer Fläche als Darstellungsebene, einem System, das die Position der Hand erkennt und einem Grafikgenerator. Die Darstellungsebene ist bei IO fest.

Der Forschungsbereich interaktive Oberflächen überschneidet sich mit vielen anderen Themengebieten. Interaktive Oberflächen werden häufig mit Tangible User Interfaces



Abbildung 1. Passanten bedienen eine IO in Form eines Schaufensters. Abbildung aus [1].

(Benutzerschnittstellen, die mit greifbaren Objekten funktionieren) und “Hands-free” Computing eingesetzt (siehe Abschnitt 10.2). Der Benutzer muss keine speziellen Eingabegeräte bedienen um mit dem Rechner zu interagieren. Zur Interaktion wird nur die Hand eingesetzt. PCs gehören dazu nicht, da sie mit Tastatur und Maus bedient werden.

3 Motivation für interaktive Oberflächen & “Table Top” Benutzerschnittstellen

Der Haupteinsatzbereich interaktiver Oberflächen ist die Unterstützung kollaborativer Arbeit in Arbeitsgruppen.

Das Einrichten einer Arbeitsgruppe hat zum Ziel ein Ergebnis zu erreichen. Dazu muß auf optimale Art und Weise miteinander kommuniziert werden können. Mit der Entstehung der digitalen Nachrichtentechnik und des Auftretens des Computers stehen hierfür sehr mächtige Werkzeuge zur Verfügung. Im Gegenzug ist jedoch auch die Komplexität der Nutzung dieser Techniken deutlich gestiegen.

Ein Beispiel hierfür: Vor 25 Jahren wurde ein Handout für einen Vortrag mit der Schreibmaschine getippt. Grafiken für den Vortrag wurden von Hand erstellt. Es wurden Folien von Hand layoutet und für die Präsentation belichtet. Zum Vortrag kam die DozentIn mit ihrem Foliensatz und dem Handout. Jeder Anwesende bekam ein Handout und der Vortrag konnte mit dem Auflegen der ersten Folie auf den Overhead-Projektor beginnen. Im Anschluß an den Vortrag wurde Gegenvorschläge, Kritikpunkte diskutiert.

Vielfach kam es gerade bei tiefergreifenden Diskussionen zu Missverständnissen und unterschiedlichen Verständnissen der Materie, die unaufgedeckt blieben.

Bis heute arbeiten Arbeitsgruppen nach dem gleichen Schema, die neuen technischen Möglichkeiten werden dem klassischen Vortrag übergestülpt. Im Ergebnis kann durch die technische Komplexität die Qualität der Kommunikation sogar sinken. Dazu ein paar Beispiele:

- Geringere Auflösung und Lichtstärke von LCD-Projektoren im Vergleich zu Dia und Overhead-Projektoren.
- Teilweise schlechte Bildqualität der VGA-Anschlüsse von Notebooks.

Es soll aber nicht unerwähnt bleiben, daß Animationen, die Einbindung von Ton und Video helfen können die Inhalte besser verständlich zu machen.

Festzustellen bleibt, daß die neuen technischen Möglichkeiten für die Gruppenkommunikation nicht voll ausgenutzt werden und daß physische Objekte unkomplizierter zu Handhaben sind (Der Foliensatz wird nicht abstürzen).

Interaktive Oberflächen können die Kommunikation innerhalb einer Gruppe wesentlich verbessern, da jeder den Inhalt ändern kann und somit iterativ gemeinsam ein Ergebnis erarbeitet wird. Mißverständnisse werden schnell offensichtlich, da durch Manipulation der IO für die anderen Teilnehmer schnell klar wird, ob man die Probleme und deren Lösung ähnlich sieht.

Darüber hinaus ist bekannt, daß der Lerneffekt beim Involvieren der Lernenden im Vergleich zur visuellen und mündlichen Wissensvermittlung wesentlich größer ist. Das heisst interaktive Oberflächen bieten auch erweiterte Möglichkeiten der Wissensvermittlung (siehe [2], S.1).

In [3] S.4 wird als Motivation für IO folgendes Szenario angegeben. Es wird die Arbeit einer kleinen Arbeitsgruppe mit Laptops dargestellt. Jeder Benutzer sieht auf seinen Bildschirm. Andere Benutzer können evtl. noch Teile davon sehen. Gemeinsames Arbeiten z.B. an einem Vortrag ist so aber schlecht möglich. Durch Verwendung des UbiTable ist dieses Problem gelöst indem die Ein- und Ausgabe auf der Tischoberfläche stattfindet (Zur Erläuterung siehe Abbildung 2). Diese ist für jeden Benutzer einsehbar. Ziel des Projektes ist es Daten, die auf mitgebrachten Laptops gespeichert sind über ad-hoc Netzwerke¹ auf dem UbiTable darzustellen und in der Gruppe zu manipulieren.

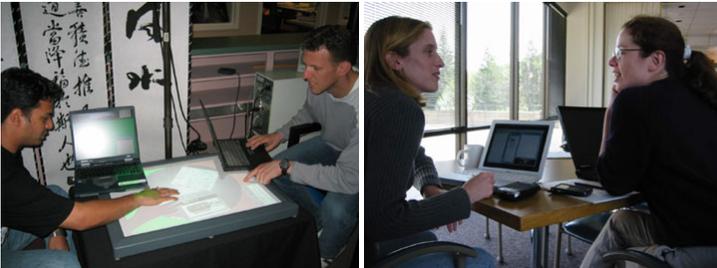


Abbildung 2. Vergleich zwischen der Zusammenarbeit am UbiTable und ohne besondere Hilfsmittel. Abbildungen aus [3] S.4.

¹ Für Ad-hoc Netzwerke ist keine besondere Konfiguration der Software notwendig damit Geräte miteinander kommunizieren können. Bluetooth und ZeroConf (<http://www.zeroconf.org>) für IP-Netze sind Beispiele hierfür.

4 Anwendungsbereiche

Da als Projektionsfläche z.B. eine feste Mauer verwendet und der Projektor an der Decke montiert werden kann, bieten sich interaktive Oberflächen als Informationssysteme in öffentlichen Bereichen, die Vandalismus-gefährdet sind an.

Ein weiteres Einsatzgebiet ist die Unterstützung von grösseren Arbeitsgruppen. Beispielsweise können bei der Städteplanung [4] alle beteiligten Gruppen direkt am Prozess teilnehmen. Jeder kann Strassen verschieben, Gebäudehöhen festlegen. Der Rest der Beteiligten kann die Änderungen sofort wahrnehmen und kommentieren. Diese Änderungen können aufgezeichnet werden. Es kann mitprotokolliert werden, wer welche Veränderungen vorgenommen hat. Dieses Szenario kann also dazu beitragen Konsens herzustellen und nachzuvollziehen von wem welche Entscheidung getroffen wurden.

Ein weiterer Anwendungsbereich sind Spiele. Am MIT wurde ein Spiel entwickelt [5], bei dem zwei räumlich getrennte Spieler über eine IO interagieren. Die Oberfläche ist durch virtuelle Kacheln abgedeckt. Hinter den Kacheln ist der andere Spieler zu sehen. Mit einem Fussball wird auf die interaktive Oberfläche geschossen. Die Intensität und die Position des Aufpralls wird mittels Mikrofonen gemessen (siehe Abschnitt 6.6). Übersteigt die Intensität in dem Areal der Kachel einen festgelegten Schwellenwert, so "bricht" sie. Ziel des Spiel ist es alle Kacheln des Gegners zu zerschliessen.

5 Vorteile gegenüber anderer Techniken

Bei manchen Projekten ist ein Einsatz von Head Mounted Displays (HMD)² oder interaktiven Oberflächen möglich. Interaktive Oberflächen bieten in diesem Fall einige Vorteile. Das Blickfeld und die Auflösung eines Head Mounted Displays ist eingeschränkt. Da bei HMDs oft noch der Rechner mitgetragen muss kann der Einsatz physisch sehr anstrengend werden [6].

Bei der Verwendung von HMDs kann ich das Gesicht, die Mimik anderer Personen nicht erkennen. Für kollaboratives Arbeiten sind HMDs daher nicht geeignet.

Die Arbeit einer Arbeitsgruppe wird sehr viel besser unterstützt, als durch andere Technologien. Parallele Eingabe und Ausgabe durch mehrere Personen ist möglich.

6 Erkennung der Hand

IO benutzen zur Interaktion keine Tastatur und Computermaus. Vielfach wird die IO durch die Hand oder den Arm direkt manipuliert [7]. Die Erkennung der Position der Hand muss in Echtzeit erfolgen. Will man mit der Hand einen virtuellen Gegenstand der IO manipulieren, dann muß es sofort darauf reagieren.

² Das Bild der grafischen Ausgabe wird direkt in die Augen projiziert. Dazu wird ein Helm verwendet. Direkt vor den Augen liegen 2 kleine LC-Displays.

6.1 Optische Erkennung

Für die optische Erkennung wird ein Bild oder mehrere Bilder gleichzeitig aus unterschiedlichen Perspektiven aufgenommen. Zusätzlich können Bilder aus unterschiedlichen Frequenzbereichen aufgenommen werden (z.B. Infrarot). Mittels Bildverarbeitung wird aus dem Bildstrom die Position der Hand ermittelt (siehe auch [8]). “Vereinfachend” ist in dieser Anwendung, daß bekannt ist, wie die projizierte Fläche aussieht. Sie entspricht dem projizierten Bild. Mittels Differenzbildung kann so aus dem Kamera-Bild die Hand herausgerechnet werden. Weiterhin ist der Abstand zwischen Kamera und Hand fast konstant. Durch diese Einschränkungen lassen sich Fehler reduzieren, die Genauigkeit erhöhen und die Verarbeitung beschleunigen (siehe Abbildung 3).

Bei Verwendung von optischen Markern³ muss der Abstand und der Winkel der

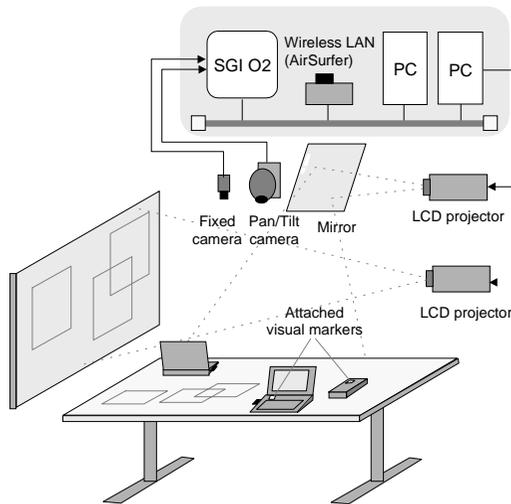


Abbildung 3. Aufbau einer IO, die mit optischer Erkennung arbeitet. Abbildungen aus [9].

zwischen Kamera und Marker auftritt berücksichtigt werden. Die Kamera muss die Charakteristika eines Markers noch erkennen können (siehe Abbildung 4). Dazu muss die Auflösung der Kamera und die Unterschiede zwischen Schwarz und Weiss groß genug sein. Als günstig hat sich die Positionierung der Kamera(s) an der Decke mit Blick auf den Tisch herausgestellt. Marker sind dann in der horizontalen Ebene anzubringen. Da das Licht von oben kommt ändert sich der Weiss- und Schwarzwert⁴

³ Marker sind klar erkennbare, orientierte Muster i.A. aus schwarzen und weißen Blöcken.

⁴ Schwarzwert/Weisswert: Intensitätswert (beispielsweise in einem CCD), der als Schwarz oder Weiss interpretiert wird. Messwerte, die darunter oder darüber liegen werden uniform als Schwarz bzw. Weiss festgelegt. Damit wird der Helligkeitsumfang festgelegt.

nur geringfügig und die geometrische Verzerrung ist gering. Vgl. auch [9].

Ein Vorteil der Bildverarbeitung ist die Möglichkeit viele Objekte gleichzeitig



Abbildung 4. Markierter Computer und Marker im Detail. Abbildungen aus [9].

erfassen zu können. Bei markierten Objekten ist eine Unterscheidung leicht. Werden unterschiedliche Hände erfasst ist es unmöglich ohne Verwendung eines Markers zu wissen zu wem sie gehört.

Problematisch bei der optischen Erkennung ist die Empfindlichkeit gegenüber sich verändernden Lichtverhältnissen. Durch den großen Rechenaufwand ist eine Echtzeitverarbeitung besonders schwer. In der Bildverarbeitung können nur Merkmale extrahiert werden, die die Kamera auflösen kann. Im Vergleich zu anderen Verfahren (z.B. kapazitive Erkennung) ist die Auflösung der Kamera eher grob.

Digital Vision Touch Technologie (DVIT) Digital Vision Touch Technologie (sowie wie Digitale Wahrnehmung der Berührung) ist eine kommerzielle Neuentwicklung im Bereich der optischen Erkennung von SMARTTech Inc. [10]. Anders als in 6.1 werden 4 Kameras in die Ecken der IO platziert. Mittels Bildverarbeitung kann aus dem Datenstrom die Position und der Abstand zur IO ein oder mehrerer Hände bestimmt werden. Sind die Finger noch zusätzlich markiert können auch Nutzer unterschieden werden.

6.2 Erkennung mittels Infrarot

Dieses Erkennungsverfahren ist für rechteckige Oberflächen geeignet. Dabei werden an den Rändern der interaktiven Oberfläche Infrarot-LEDs angebracht. An den gegenüberliegenden Seite wird mittels einer Photodiode das Licht aufgenommen. Aus dem fehlenden Licht an 2 Photodioden kann die Position eines Objektes bestimmt werden. Bei der Benutzung mit mehreren Händen kann es zu Verdeckungen kommen. Es

lässt sich nur die Position eines oder mehrere Objekte bestimmen; es kann nicht erkannt werden um welches Objekt es sich handelt. Die Größe einer LED ist nicht beliebig klein, dadurch ist die Auflösung dieses Verfahrens eher gering [10]. Diese Art der Erkennung ist ziemlich eingeschränkt nutzbar.

6.3 Erkennung mittels Ultraschall

Um die Position eines Objektes im 3-dimensionalen Raum bestimmen zu können sind mindestens drei Ultraschallemitter notwendig. Das zu trackende Objekt besitzt einen Ultraschallempfänger. Der Ultraschallempfänger kann die drei Emitter unterscheiden. Ist die Position der drei Sender bekannt kann die Position des Empfängers aus den Laufzeiten der Signale der 3 Emitter errechnet werden.

Vorteile des Erkennung mittels Ultraschall sind geringe Latenz der Positionsbestimmung, Genauigkeit unter konstanten Randbedingungen, geringe Größe der Empfangskombination, daß Sender und Empfänger sich optisch nicht "sehen" müssen und daß viele Objekte gleichzeitig beobachtet werden können.

Nachteile sind die relativ kleine Erfassungsbereich und die Empfindlichkeit gegenüber Randbedingungen wie Luftfeuchtigkeit und -temperatur. Berücksichtigt werden muss, daß Ultraschall Echos erzeugen kann. Zum vorhergehenden siehe [11].

Für interaktive Oberflächen haben diese Faktoren aber keine oder nur geringe Bedeutung.

6.4 Erkennung mittels Lasermessung

Aus einer Ecke der Projektionsebene wird ein rotierender Laserstrahl ausgesendet. Dieser deckt die Projektionsebene vollständig ab. Bei Kontakt mit der Hand wird ein Teil des Laserstrahls zurückreflektiert. Aus dem Winkel der Reflexion und dem Anteil des Signals, das reflektiert wird lässt sich die Position der Hand bestimmen [12].

Es ist günstig einen Laser im sichtbaren Spektrum zu verwenden. Der Benutzer kann dann sehen, wann er im sensitiven Bereich ist. Bei der Benutzung mit nur einer Hand kann die Position sehr genau bestimmt werden. Kommen mehrere Hände ins Spiel kann es zu Verdeckungen kommen.

6.5 Kapazitive Erkennung

Dieses Verfahren beruht auf elektrischen Gesetzen. Unter der empfindlichen Oberfläche liegt ein Geflecht leitender Drähte. Die Hälfte der Drähte sind parallel zueinander die anderen im 90 Grad Winkel dazu. Die Drähte berühren sich aber nicht. Auf die parallelen Drähte wird jeweils ein eindeutiges Referenzsignal (z.B. phasenverschobene Sinuswellen) gelegt. An den Kreuzungspunkten kommt es zwischen den Drähten zu kapazitiven Effekten. Das heisst die in 90 Grad liegenden Drähte sind kapazitiv gekoppelt. Bewegt sich ein leitfähiges Objekt, wie eine Hand über dem Drahtgitter, dann ändert sich das gekoppelte Signal. Aus der Änderung kann die Position der Hand bestimmt werden (siehe [7], S.2). Zur Funktionsweise siehe auch Abbildung 5.

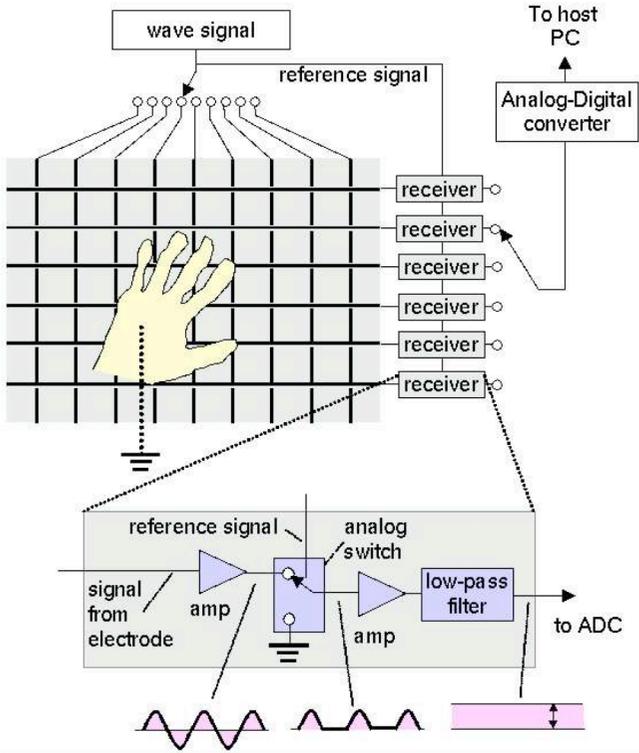


Abbildung 5. Physikalischer Aufbau der SmartSkin. Abbildung aus [7]

Die Position der Hand wird sofort erfasst, es gibt keine Latenz. Die Genauigkeit des Systems lässt sich durch ein engeres Drahtgitter erhöhen. Es können mehrere Hände gleichzeitig erfasst werden.

DiamondTouch In [13] wird von Paul Dietz und Darren Leigh DiamondTouch vorgestellt. Dieses System verwendet auch kapazitive Kopplung. Im Gegensatz zu SmartSkin wird für jeden "Punkt" der IO ein Erkennungselement verwendet. Als Erkennungselement werden Antennen, die mit unterschiedlich-frequenten Strom belegt sind eingesetzt. Am Stuhl des Benutzers ist ein Empfangsgerät angeschlossen. Nähert sich ein Finger einer Antenne, so beeinflusst das elektromagnetische Feld die Frequenz der Schwingung in Antenne und Empfangsgerät. Aus dem neuen Signal lässt sich bestimmen welcher Sender und Empfänger kapazitiv gekoppelt sind oder anders ausgedrückt welcher Benutzer gerade welchen Punkt der IO berührt. Über die Stärke der kapazitiven Kopplung kann ein Abstandsmas definiert werden. Es ist gut möglich, daß ein Finger mehrere Antennen berührt. Auch dieser Fall stellt kein Problem da, da der Hauptberührungspunkt am stärksten im erzeugten Signal herauskommt. Der Durchmesser einer Antenne ist im Moment 0,5 cm groß. Die Auflösung kann mittels Interpolation auf 0,1 mm erhöht werden. Siehe Abbildung 6

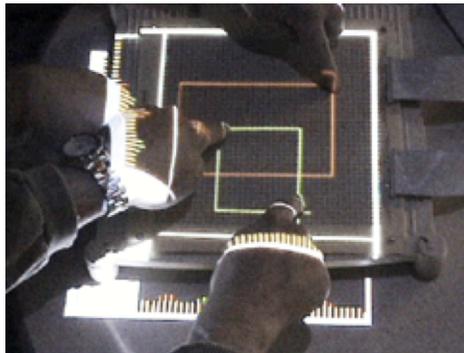


Abbildung 6. Zwei Benutzer bei der Benutzung von DiamondTouch [13]. An den unterschiedlichen Farben ist gut zu erkennen, daß zwei Benutzer unterschieden werden können. Dem System ist es zusätzlich möglich mehrere Kontaktpunkte gleichzeitig zu erkennen.

6.6 Erkennung über Schall

Klopft man auf ein Material breiten sich die Schallwellen ringförmig von der Klopfstelle aus. Ist die Größe der Ebene bekannt, kann über an den 4 Ecken positionierten Mikrofonen die Ursprungsposition und die Intensität des Schalls bestimmt werden (siehe [12]).

Eine Figur mittels dieser Art der Erkennung zu zeichnen ist unmöglich. Für Dialog orientierte Systeme, wie die in Abschnitt 1 vorgestellte Autokonfigurations-Anwendung ist die Erkennung über Schall sehr gut geeignet.

Diese Art der Erkennung ist auch besonders für Vandalismus-gefährdete oder dem Wetter ausgesetzte Installationen geeignet. Die Mikrophone sind hinter (einer dicken Glasscheibe) angebracht und das Projektionssystem kann an einer sicheren Stelle, abgekapstelt angebracht werden [1] S.3.

6.7 Zusammenfassung der Erkennungsverfahren:

Zusätzlich zur Erkennung der Position der Hand und Orientierung greifbarer Objekte werden in [13] folgende Kriterien für die Güte einer Erkennungstechnologie genannt: Mehrere Berührungen können gleichzeitig erkannt werden; Der Kontaktpunkt kann einer Person zugeordnet werden; Objekte außerhalb der IO interferieren nicht mit der Erkennung; Die Erkennung ist unempfindlich gegenüber Umwelteinflüssen; Die Finger der Hand reichen zur Benutzung aus. Billige Herstellung ist möglich. Jun Rekimoto nennt in [7] zusätzlich die Erkennung des Abstands bzw. des Drucks auf die IO.

Warum sind diese Merkmale wichtig? Die Position der Hand muss erfasst werden, damit ich weiß, wo der Benutzer gerade mit der IO interagiert. Kenne ich die Position so kann ich die Ausgabe darauf anpassen und beispielsweise ein Menü darstellen. Die Orientierung greifbarer Objekte festzustellen ist wichtig um sie richtig in die IO einzubinden (siehe Abschnitt 10.2). Ich muss gleichzeitig Kontaktpunkte erkennen, damit mehrere Finger oder Personen gleichzeitig auf der IO arbeiten können. Die Zuordnung ein oder mehrerer Kontaktpunkte zu einer Person erlauben es nachzuvollziehen wer was geändert hat und ermöglicht Sicherheitsmechanismen. Die Erkennung einzelner Finger ist wünschenswert damit auch kleinere Details auf einer IO geändert werden können und ermöglichen neue, schnelle Interaktionsmuster (siehe auch Abschnitt 10.1). Durch einen Schwellenabstand/-Druck kann ein Moduswechsel ausgelöst werden oder beispielsweise kann beim Zeichnen einer Linie die Dicke über den Druck variiert werden. Die optische Erkennung, Erkennung mittels Ultraschall, Infrarot und Lasermessung können die Position der Hand oder anderer Objekte bestimmen, sie können aber nicht ohne Einbeziehung von Hilfskonstrukten (Klicker, Gesten, Festlegen zweier Interaktionsebenen) erkennen, ob ein bestimmter Modus aktiv ist. Bei der Ermittlung der Position mittels Schall, kapazitiver Kopplung oder DiVT kann ein Moduswechsell durch Erhöhen des Drucks oder Verringern des Abstands erkannt werden. Diese Form der Interaktion ist sehr natürlich und somit intuitiv. Wenn ich auf Papier einen Strich zeichnen will muss ich das Papier mit dem Stift berühren. Der Druck legt dann fest, wie der Strich aussieht. Diese Analogie lässt sich sehr gut auf interaktive Oberflächen übertragen.

Aufgrund der horizontalen Orientierung der IO werden oft auf ihr greifbare Objekte (tangible bits) eingesetzt. Erkennungsverfahren, die nur die Position eines Objektes erfassen können sind daher für den Einsatz auf IO ungeeignet. Für diese Objekte kann eine andere Erkennung als die für die Hand eingesetzt werden (z.B. Ultraschall). Im Bereich der Erweiterten Realität wird das schon praktiziert.

7 Projektionssystem

In diesem Paper werden nur feststehende, interaktive Oberflächen betrachtet. Es soll aber nicht unerwähnt bleiben, daß es Systeme gibt, in denen sich Projektionsflächen ändern. In diesem Fall muss dann die Beschaffenheit der Oberfläche berücksichtigt werden auf die projiziert wird.

Da auf IO potenziell sehr informationsdichte Daten (z.B. Karten) dargestellt werden, sollte die Auflösung des Monitors oder des Projektors sehr hoch sein.

7.1 Selbstleuchtende Oberflächen (Monitore)

Monitore werden hauptsächlich in "Table-Top" Systemen eingesetzt. Dabei sitzt der Monitor in einer Halterung und die Bildfläche zeigt nach oben.

Vorteil dieser Lösung ist, daß Monitore aufgrund der hohen Leuchtkraft auch draussen eingesetzt werden können. Da ein Monitor aber eine festgelegte Bildschirmdiagonale hat, sind damit nicht beliebig große Flächen realisierbar. Dafür werden Projektoren verwendet.

7.2 Projektionsbasierte Darstellung

Für die Projektion werden handelsübliche LCD/DLP-Projektoren eingesetzt. Sie besitzen inzwischen genügend Leuchtkraft und Kontrast, um in nicht komplett abgedunkelten Räumen verwendet werden zu müssen. Durch "Parallelschaltung" lassen sich beliebig große Flächen ausleuchten.

Grundsätzlich kann das Bild von vorne oder von hinten auf die IO projiziert werden. Allerdings lässt sich die Rückprojektion nicht mit jeder Erkennungsmethode verwenden. Ein Beispiel dafür ist die kapazitive Erkennung. Die Entscheidung für Rück- oder Frontprojektion ist wichtig. Werden auf der interaktiven Oberfläche greifbare Objekte eingesetzt und soll auf diese Objekte Grafik projiziert werden (beispielsweise, welche Musikspur gerade manipuliert wird [14]) dann muss die Frontprojektion eingesetzt werden.

Die meisten Forschungsprojekte verwenden zur Darstellung LCD/DLP-Projektoren und Frontprojektion.

8 Bisherige Veröffentlichungen

In diesem Abschnitt werden weitere Arbeiten im Bereich interaktiver Oberflächen vorgestellt.

8.1 Erweiterungen von Whiteboards:

Whiteboards sind abwaschbare, hängende Oberflächen, auf denen mit speziellen Stiften geschrieben werden kann. Das Geschriebene ist permanent. Durch Verwendung eines speziellen Schwamms können Inhalte gelöscht werden. Das Whiteboard funktioniert also wie eine Tafel. Durch beispielsweise kapazitive Erkennung kann die Position von

Stiften auf dem Whiteboard bestimmt werden. Diese Entwicklung führte zur Entwicklung der Smartboards [15].

Dabei wird nicht mehr mit echter Tinte auf die weiße Oberfläche, sondern durch beobachtete Objekte (tracked objects) virtuell auf die Oberfläche gezeichnet. Ein beobachtetes Objekt kann in diesem Fall ein Griffel sein, der malt und ein Quader der Inhalte löscht. Die Projektion wird dabei beispielsweise von einem LCD-Projektor übernommen. Dadurch lassen sich bestehende Computeroberflächen einbinden. Es ist zum Beispiel möglich eine Vorlesung mit PowerPoint zu halten und währenddessen mit dem interaktiven Stift Anmerkungen und Zeichnungen hinzuzufügen. Diese Daten können aufgezeichnet werden und den Studenten zur Verfügung gestellt werden. Die Vorlesung kann so deutlich flexibler auf Anregungen und Fragen aus dem Plenum eingehen und das gemeinsam Erarbeitete bleibt erhalten.

XEROX hat sich mit der Frage beschäftigt welche Interaktionsmöglichkeiten durch die neue Technik möglich werden. In [16] wird die Verwendung eines erweiterten Smartboards im Büro untersucht. Beispielsweise wird dem Benutzer immer Platz zur Eingabe zur Verfügung gestellt. Bestehende Inhalte werden verschoben und verkleinert, um Platz frei zu machen. (siehe [16], S.3 und Abbildung 7). Das Smartboard stellt dem Benutzer Seiten zur Verfügung. Will man eine freie Fläche, "blättern" man um, indem mit dem Stift am unteren linken oder rechten Ende "gezogen" wird. Dieses Verhalten ist analog zur Bedienung einer Flipchart (siehe [16], S.4). Ein Beispiel zur Verwertung der Inhalte des Smartboards ist der eingebaute Rechner. Auf dem Smartboard werden

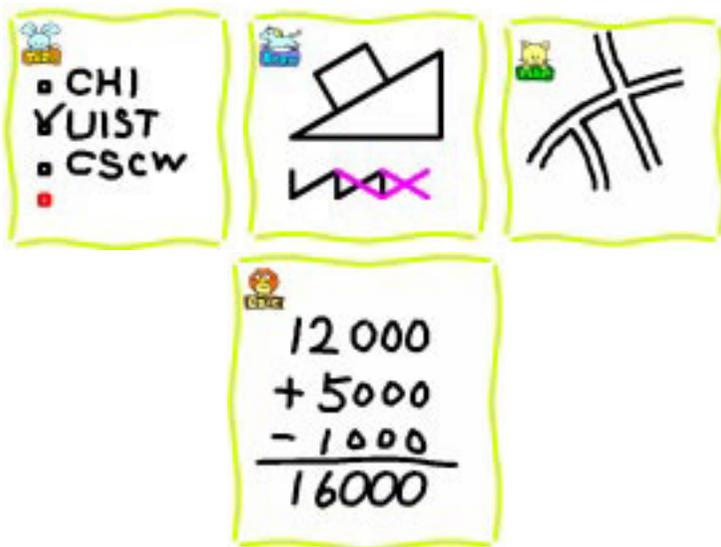


Abbildung 7. Verschiedene Arbeitsmodi auf dem Smartboard Flatlands von XEROX Parc [16]. Der Reihenfolge nach Auswahl, Zeichnung, Kartenerstellung und Rechnen

Additionen und Subtraktionen handschriftlich untereinander geschrieben. Über OCR⁵ werden die Summanden erkannt. Unter der Rechnung wird der Endbetrag eingeblendet. Dazu wird ein handschriftlicher Stil verwendet, um den informellen Charakter eines Whiteboards zu erhalten.

Die Inhalte des Whiteboards können gesichert und wiederhergestellt werden. Dazu werden keine Dateinamen verwendet, sondern der Benutzer gibt einen Zeitraum, Inhalt oder andere Metadaten an. Auf eine Anfrage werden vom System die Treffer als Thumbnails⁶ dargestellt.

8.2 Computer-unterstütztes, kooperatives Arbeiten (CSCW)

Im Mitsubishi Electronic Research Laboratory haben sich Shen, et al darüber Gedanken gemacht, wie sich IO für die Arbeit in (kleinen) Arbeitsgruppen einsetzen lassen [3]. Es wurden spontane Meetings beobachtet. Dabei stellte sich heraus, daß horizontale Oberflächen genutzt wurden um darauf Unterlagen zu legen und sie zu bearbeiten. Die Orientierung des Dokuments legt dabei fest, ob es privat oder öffentlich ist. Wird ein Dokument in die Mitte gelegt wird dadurch impliziert, daß ein Kommentar der anderen Gruppenmitglieder erwünscht ist. Ist hingegen das Dokument an der Kante des Tisches und nur zu einer Person orientiert so ist implizit klar, daß der Inhalt "privat" ist. In Abschnitt 3 wurde bereits erläutert, daß bei der Arbeit mit Laptops diese Art des sozialen Protokolls nicht mehr umgesetzt werden kann. Die zu bearbeitenden Dokumente bezieht der UbiTable von einem verbundenen Laptop. Für n Nutzer gibt es n private Bereiche an den Seiten des Tisches und in der Mitte des Tisches einen gemeinsam genutzten Bereich. Die kognitive Belastung ist bei der Bedienung des UbiTable sehr gering, da z.B. das "Senden" eines Dokuments an einen anderen Teilnehmer der Arbeitsgruppe durch ein einfaches Draggen in dessen privaten Bereich möglich ist. Einfacher lässt sich diese Aktion nicht mehr gestalten. Diese Aussage wurde durch [17] bestätigt. Daß dies ein wichtiger Faktor in der Arbeitswelt ist stellt Martina Schütze in einer Studie da [18]. Dabei wurden Maschinenbau-Studenten gebeten einen neuen Grill zu entwerfen. Die eine Hälfte der Gruppe verwendete nur Papier und Stift zum Entwurf, die andere verwendete ein CAD-Programm⁷ dazu. Die Ergebnisse wurden von einer Experten-Gruppe beurteilt mit dem Ergebnis, daß sich in der Gruppe, die nur mit Papier und Stift gearbeitet haben die kreativeren Resultate fanden. Die Studie legt nahe, daß die Unterschiede nur aufgrund der zusätzlichen kognitiven Belastung durch die Nutzung des CAD-Programms auf einem PC zustande gekommen sind.

⁵ Optical Character Recognition (OCR): Optische Schriftzeichen-Erkennung. Software, die gescannte Textdokumente oder Handschrift erkennen kann.

⁶ Verkleinerte Darstellung in der Größe eines Daumennagels.

⁷ CAD steht für Computer Aided Design. Mit CAD-Programmen können am Rechner hoch komplexe geometrische Modelle erstellt werden, die dann wiederum von anderen Programmen direkt z.B. für Simulationen und prototypische Herstellung genutzt werden können.

9 Grafikerzeugung

9.1 Was soll dargestellt werden?

I.A. werden zur Benutzung einer grafischen Oberfläche Menü, Kontextmenü, Werkzeugauswahl und Moduswechsel gebraucht. Für das Einbringen dieser Merkmale in IO gibt es viele Möglichkeiten. Aus diesen Möglichkeiten hat sich noch kein Königsweg herauskristalisiert. Auch deshalb, da die Oberflächen unterschiedlich groß sind und von unterschiedlich großen Gruppen verwendet werden [19]. Gesetzmäßigkeiten wie Fitt's Law gibt es noch nicht.

Wenn bekannt ist welcher Nutzer sich wo gerade befindet (vgl. 6.5, 6.1) so kann in seiner Nähe sein (privater) Arbeitsbereich angezeigt werden. Jeder private Arbeitsbereich soll von den anderen über graphische Merkmale abgegrenzt werden. Dazu bieten sich unterschiedliche Texturen und Farben an.

Der Bereich der IO, der für die gemeinsame Arbeit verwendet wird kann nie perfekt auf eine am Tisch verteilte Gruppe ausgerichtet werden. Um dieses Problem zu mindern kann dieser Bereich rotiert werden oder es wird eine zweite Darstellung an die Wand projiziert [19].

Da IO vor allem in Gruppen deren Mitglieder um einen Tisch versammelt sind verwendet werden, ist unklar wo und zu wem orientiert ein Menü angezeigt werden soll. Besser ist es in dieser Anwendung Kontextmenüs zu verwenden, die durch die Hand oder einzelne Finger aufgerufen werden (siehe Abbildung 8). Ansonsten bieten sich Icons an, da sie besser erkannt werden können, als eine potenziell zu kleine und falsch orientierte Schrift.

Zusätzlichen Nutzen gegenüber nicht Computer-unterstützten Installationen bieten IO durch Erzeugung interaktiver Grafiken. In der Städteplanung wird über interaktive Oberflächen Dynamik eingebracht. Es ist möglich Schattenwurf in Abhängigkeit der Tageszeit und des Datums darzustellen. Das Modell kann zum Leben erweckt werden, in dem der Personenverkehr und Fahrzeugverkehr simuliert und im Modell dargestellt wird (siehe [4], S.6).

Generell sind interaktive Oberflächen gut für Designaufgaben geeignet. So könnte das Modell eines Autos während einer Strömungssimulation verändert werden [20].

10 Interaktion

Da interaktive Oberflächen nicht mit Maus und Tastatur gesteuert werden, sondern mit der Hand und den Fingern müssen für die Interaktion völlig neue Ansätze entwickelt werden.

Es ist dabei schwierig und vielleicht auch gar nicht erwünscht auf die Paradigmen der Desktop-Welt zurück zu greifen.

Für die intuitive Bedienung IO ist es notwendig zwischen Modi unterscheiden zu können. Um zwischen Modi zu wechseln können unterschiedliche Mittel verwendet werden. So kann z.B. eine Geste zwischen Skalierung und Verschiebung in einer Architekturanwendung umschalten, ein Werkzeug könnte dann aktiv sein, wenn die Hand

Mittels kapazitiver Erkennung können nicht nur einzelne Hände sondern einzelne Finger unterschieden werden. Dadurch eröffnen sich ganz neue Interaktionsmöglichkeiten. Bei Manipulationen, bei denen zwei Werte (gleichzeitig) verändert werden, kann dies mittels zweier Finger gleichzeitig passieren. Ein Beispiel dafür ist die Manipulation zweier Ankerpunkte einer Bezierkurve oder Skalierung mit Drehung eines Bildes (siehe Abbildung 10). Desweiteren ist es möglich an jeder Fingerspitze ein Menüpunkt

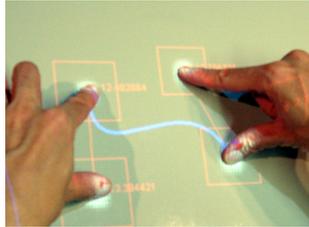


Abbildung 10. Gleichzeitige Veränderung zweier Ankerpunkte einer Bezierkurve auf einer interaktiven Oberfläche mit kapazitiver Erkennung. Abbildung aus [7].

darzustellen.

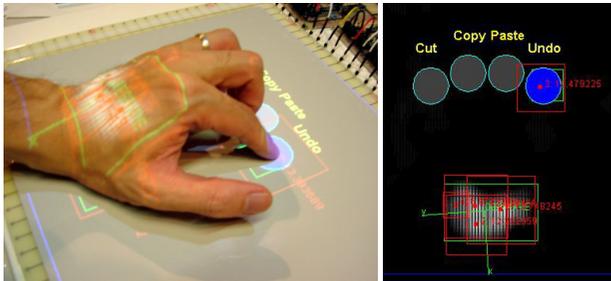


Abbildung 11. Finger-gesteuerte Menüdarstellung auf einer interaktiven Oberfläche mit kapazitiver Erkennung. Abbildungen aus [7].

10.2 Greifbare Objekte (tangible bits)

Menschen sind an die Manipulation von greifbaren Objekten gewöhnt. Die Manipulation von greifbaren Objekten ist intuitiv und natürlich. Ein Beispiel hierzu. Ein Kind kann leicht eine Kuh mit Lehm modellieren. Es wird dem Kind sogar Spass machen. Das gleiche Ergebnis in einem 3-D Modeller zu erzielen ist ungleich schwieriger und zeitaufwendiger. Daher können greifbare Objekte in einer Benutzerschnittstelle sehr

mächtige Werkzeuge sein. Im Kontext interaktiver Oberflächen können Holzklötze unterschiedlicher Form auf Werkzeuge abgebildet werden. Ist einem Klotz eine Funktion, wie Rotieren oder Skalieren zugeordnet, dann muss noch festgelegt werden, unter welchen Bedingungen das Werkzeug aktiv ist. Durch Verwendung von Gesten im Umgang mit dem Klotz (Abdeckung des Klotzes durch die Hand, Greifen des Klotzes an der langen oder kurzen Seite, Abdecken des Klotzes mit beiden Händen) wird die Funktion des Klotzes mit einem virtuellen Objekt assoziiert und im nächsten Schritt aktiviert [21]. Eine andere Möglichkeit ist Daten mit einem greifbaren Objekt zu assoziieren. Im Szenario der Städteplanung könnte man jedem Gebäude einen Klotz zuordnen. Position und Orientierung im Raum des Gebäudes werden dann direkt über den Klotz festgelegt (vgl. [22]). Es stellt sich dann die Frage, ob auf die greifbaren Objekte gar nichts oder etwas spezielles projiziert werden soll. Um die Darstellung auf dem greifbaren Objekt synchron zu halten muss die Position, die Orientierung und die Form des greifbaren Objekts bekannt sein oder erkannt werden.

Die Form des Objekts kann durch ein 3-D Modell bekannt gemacht werden. Mittels der Erkennung durch Ultraschall bzw. der optischen Erkennung lässt sich Position und Orientierung bestimmen.

Bei der optischen Erkennung werden die greifbaren Objekte mit einem oder mehreren Markern versehen. Im Bild der Kamera wird nach diesen Markern gesucht. Wird ein Marker gefunden wird aus dessen Verzerrung die Orientierung des greifbaren Objekts bestimmt. Wird ein greifbares Objekt von einem Benutzer berührt um es woanders hinzusetzen oder um es zu drehen so werden die optischen Marker eventuell von der Hand verdeckt. Die Darstellung der IO wird dann nicht mehr angepasst. Das ist ein Nachteil gerade wenn mit dem Objekt die Darstellung auf der IO gesteuert werden soll. Ein Beispiel hierfür ist ein greifbares Objekt das den Zoomfaktor der Darstellung steuert. Erst wenn das greifbare Objekt losgelassen wird stimmt auch wieder die Darstellung.

Die Erkennung mittels Ultraschall ermöglicht eine robuste Einbindung greifbarer Objekte in IO. Dazu wird in das greifbare Objekt eine feste Kombination von mindestens drei Ultraschallempfängern eingebaut. Gepaart mit mindestens drei Sendern lässt sich aus den Signallaufzeiten die Position und Orientierung der greifbaren Objekts bestimmen. Es bestehen die gleichen Vor- und Nachteile des Verfahrens wie in Abschnitt 6.3 beschrieben [6].

Jun Rekimoto arbeitet an der Erkennung von greifbaren Objekten mittels kapazitiver Erkennung [7]. Dazu wird das Objekt mit einem Leiter umwickelt. Der Leiter muß so um das Objekt gewickelt werden, daß die Leiterschlaufen eine eindeutige Form auf jeder Seite ergeben. Die Leiter erzeugen ein elektromagnetisches Feld, wenn durch sie Strom laufen kann. Dies ist dann der Fall, wenn der Benutzer sie anfasst und damit erdet. In einem Leiter auf der interaktiven Oberfläche wird dadurch eine Resonanz erzeugt. Da die Resonanz abhängig von den elektrischen Eigenschaften und der Form der Leiterbahnen des Objekts auf der Oberfläche ist, können dadurch Objekte und deren Seiten unterschieden werden. Wie gut diese Technik in der Praxis funktioniert wurde noch nicht untersucht.

10.3 Zusammenfassung der Interaktion mit IO

Interaktion mit einer IO ist dann besonders intuitiv, wenn mehrere Ein-/Ausgabe-Technologien zusammen verwendet werden z.B. greifbare Objekte auf die zusätzlich noch Grafik projiziert wird.

Bei der Größe einer IO muss berücksichtigt werden [19], daß Ein-/Ausgabe-Elemente für eine Person noch erreichbar und erkennbar sein müssen. Denkt man beispielsweise an eine Verschiebung eines Dokuments von Nutzer A zu Nutzer B durch dragging des Dokuments aus dem privaten Bereich von Nutzer A zu Nutzer B, so könnte es bei einer großen IO nötig sein mit dem Finger auf dem Dokument um den Tisch herumlaufen zu müssen.

11 Accessoires

11.1 Zusätzliche Displays

Im metaDESK Projekt [22] wird über einer interaktiven Oberfläche ein an einem schwenkbaren Arm befestigten Monitor eine andere Repräsentation der gleichen Daten unter dem gleichen Blickwinkel dargestellt (siehe Abbildung 12). Diese beiden Sichten sind gelockt. Ändert sich beispielsweise die Skalierung der Darstellung auf der interaktiven Oberfläche, dann ändert sich auch die Darstellung auf dem schwenkbaren Monitor entsprechend. Diese zusätzlichen bewegbaren Displays eignen sich besonders für

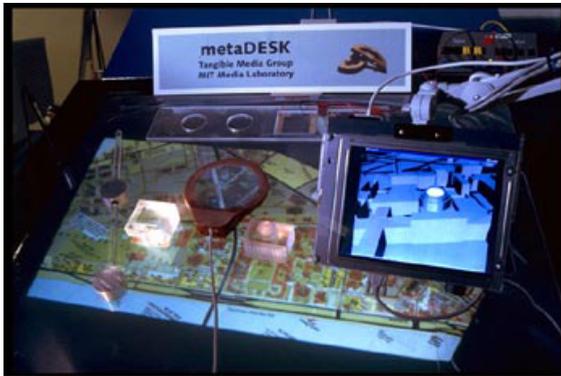


Abbildung 12. Verknüpfung zweier Darstellungen einer Stadt. Auf der IO wird eine topologische Karte gezeigt. Auf dem schwenkbaren Monitor wird eine 3-D Darstellung der Objekte, die hinter dem Monitor auf der Karte liegen gezeigt. Der Monitor verhält sich wie eine Linse. metaDesk Installation [22].

die Darstellung zusätzlicher Daten, die die Darstellung auf der interaktiven Oberfläche überfrachten würde. In der Städteplanung könnte das z.B. das Flussmodell des Windes sein, oder eine Simulation des Besucherverkehrs.

11.2 Bewegte Objekte

Mit interaktiven Oberflächen werden vielfach greifbare Objekte zur Interaktion verwendet [2]. Dabei werden über einen Gegenstand Daten der interaktiven Oberfläche manipuliert (z.B. die Position eines Gebäudes in einer Stadtsimulation). Will man in diesem Szenario beispielsweise den dargestellten Ausschnitt verändern müssen manuell alle greifbaren Objekte umgesetzt werden. Am MIT Media Lab wurde eine Technologie vorgestellt, die greifbare Objekte verschieben kann [23]. Damit ist es möglich den internen Datenstand der Anwendung und die erfassten physischen Objekte synchronisiert zu halten.

12 Zukünftige Anwendungsfälle, Ausblick

Interaktive Oberflächen werden bisher in klar definierten Konstellationen in genau definierten Anwendungsbereichen eingesetzt. Jedes Forschungssystem setzt auf andere Möglichkeiten der Eingabe und Darstellung. Es gibt nicht die von PCs bekannte Kombination von Fenstern, Maus, Tastatur und Icon. Jede interaktive Oberfläche erfordert vom Benutzer neues Verstehen und Erlernen der Interaktion. Dieser Zustand ist unerwünscht. In der Zukunft werden sich Axiome der Benutzerinteraktion interaktiver Oberflächen herauskristallisieren. Vielfach ist die grafische Darstellung hard-coded in der Implementierung zu finden. Es gibt keine bestehenden Softwarekomponenten zum Aufbau von interaktiven Oberflächen, keine GUI-Komponenten, die Events versenden etc. Eine vom Endbenutzer festlegbare Darstellung ist somit unmöglich. Es ist zu hoffen, daß es mittels Autorenwerkzeugen und Frameworks möglich sein wird, interaktive Oberflächen stärker für den Endbenutzer zu individualisieren. Für die EntwicklerInnen werden sich Softwarekomponenten und Interaktionsmuster parallel zu der Entwicklung von GUIs auf PCs herausbilden. Insgesamt werden interaktive Oberflächen in Zukunft breiter und günstiger nutzbar werden.

Literatur

1. Paradiso, J.: Tracking contact and free gesture across large interactive surfaces. Proceedings of CACM 2003 (2003)
2. Rekimoto, J., Saitoh, M.: Augmented surfaces: A spatially continuous work space for hybrid computing environments. Proceedings of SIG CHI 1999 (1999)
3. Shen, C.: Ubitable: Impromptu face-to-face collaboration on horizontal interactive surfaces. Proc. UbiComp 2003 (2003)
4. Ishii, H., Underkoffler, J., Chak, D., Piper, B.: Augmented urban planning workbench: Overlaying drawings, physical models and digital simulation. Proceedings of ISMAR 2002 (2002)
5. Mueller, F., Agamanolis, S., Picard, R.W.: Exertion interfaces for sports over a distance. Proceedings of UIST 2002 (2002)
6. Klincker, G.: Vorlesung Augmented Reality (WS 2004/2005) TU Muenchen.
7. Rekimoto, J.: Smartskin: An infrastructure for freehand manipulation on interactive surfaces. Proceedings of SIG CHI 2002 (2002)
8. Faugeras, O.: Three-Dimensional Computer Vision. The MIT Press (1993)

9. Rekimoto, J.: Multiple-computer user interfaces: Beyond the desktop direct manipulation environments. Proceedings of SIG CHI 2000 (2000)
10. Smart Technologies Inc.: Dvit digital vision touch technologie whitepaper (2004) http://www.smarttech.com/dvit/DViT_white_paper.pdf.
11. Rolland, J., Davis, L., Baillet, Y.: A survey of tracking technology for virtual environments. In: Fundamentals of Wearable Computers and Augmented Reality. (2001) 67–112
12. Paradiso, J., et al: Sensor systems for interactive surfaces. IBM Systems Journal **39** (2000) 892–914
13. Dietz, P., Leigh, D.: Diamondtouch: A multi-user touch technology. Proceedings of UIST 2001 (2001)
14. Patten, J., Retch, B., Ishii, H.: Audiopad: A tag-based interface for musical performance (2002)
15. Smart Technologies Inc.: Smartboards (2005) <http://www.smarttech.com>.
16. Mynatt, E.D., Igarashi, T., Edwards, W.K., LaMarca, A.: Flatland: New dimensions in office whiteboards. In: CHI. (1999) 346–353
17. Ringel, M., Ryall, K., Shen, C., Forlines, C., Vernier, F.: Release, relocate, reorient, resize: Fluid techniques for document sharing on multi-user interactive tables. Proceedings of SIG CHI 2004 (2004)
18. GEO: Weniger Phantasie vor dem PC (2004) http://www.geo.de/GEO/wissenschaft_natur/technik/2004_07_GEOskop_cad/?linkref=geode_teaser_toc_text&SDSID=abrufbar.
19. Ryall, K., Forlines, C., Shen, C., Morris, M.R.: Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. Proceedings of the Conference of Computer Supported Cooperative Work 2004 (2004)
20. Westermann: Vorlesung Computational Visualisation (WS 2004/2005) TU Muenchen.
21. Rauterberg, M., Fjeld, M., Krueger, H., Bichsel, M., Leonhardt, U., Meier, M.: Build-it: A planning tool for construction and design. Proceedings of SIG CHI 1998 (1998)
22. Ullmer, B., Ishii, H.: The metadesk: Models and prototypes for tangible user interfaces. In: ACM Symposium on User Interface Software and Technology. (1997) 223–232
23. Pangaro, G., Maynes-Aminzade, D., Ishii, H.: The actuated workbench: Computer-controlled actuation in tabletop tangible interfaces. Proceedings of UIST 2002 (2002)

Projizierte Benutzerschnittstellen und Steuerbare Projektoren

Matthias Schicker

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
schicke@informatik.uni-muenchen.de

Zusammenfassung Die folgende Arbeit gibt einen Überblick über Benutzerschnittstellen, die durch einen Projektor dynamisch auf verschiedenen Oberflächen erzeugt werden. Dabei liegt der Schwerpunkt auf Technologien, die keine weiteren technischen Maßnahmen außer dem Projektor (möglicherweise im Verbund mit einer Kamera) benötigen; insbesondere werden Systeme, die auf berührungsempfindlichen Oberflächen (TouchScreens) basieren oder bei denen der Nutzer herkömmliche oder spezielle Eingabegeräte (wie Datenhandschuhe, Stifte, o.ä.) verwenden muss, nur untergeordnet betrachtet. Im Wesentlichen werden zwei Geräteklassen vorgestellt: einerseits festinstallierte, schwenkbare Projektor-Kamera-Einheiten, welche beliebige Oberflächen des Raums, in dem sie installiert sind, als Projektionsfläche verwenden, um herkömmliche Bildschirmflächen zu ersetzen oder zusätzliche Informationen zu vorhandenen Objekten anzuzeigen; andererseits kleine, tragbare Projektoren im Verbund mit Kamera und ggf. weiteren Sensoren. Interaktion erfolgt bei letztgenanntem durch Schwenken des Gerätes und Benutzung von auf diesem angebrachten Bedienelementen. Zu beiden Kategorien werden Prototypen, die wichtigsten Softwaretechniken (Bildkorrektur, Tracking, etc.) und Anwendungsbereiche vorgestellt in der Reihenfolge möglichst aufeinander aufbauender Technologien. Die Arbeit erhebt keinen Anspruch auf Vollständigkeit; insbesondere die vorgestellten Prototypen sind als exemplarisch aufzufassen.

1 Einleitung

Digitale Projektoren (im Folgenden gelegentlich als „DP“ abgekürzt) sind traditionell statische Geräte, die hauptsächlich in Präsentationen als Ersatz herkömmlicher Dia- und Overhead-Projektoren eingesetzt werden. Für andere Zwecke scheinen sie in ihrer früheren Erscheinungsform auch kaum geeignet: Projektoren waren lange Zeit zu groß und schwer, um wirklich portabel zu sein, und zu teuer, um sie an Einzelarbeitsplätzen zu verwenden. Zusätzlich war erheblicher Kalibrierungsaufwand und - auf Grund geringer Leuchtstärke - eine spezielle Leinwand notwendig, damit überhaupt akzeptable Bildqualität erreicht werden konnte. Durch stete Weiterentwicklung und neue Technologien hat sich dieses Bild in den letzten Jahren allerdings zunehmend gewandelt: Projektoren der neuesten Generation sind klein, billig, leicht-

stark und erreichen bisher Monitoren vorbehaltene Auflösungen. So zeichnen sie die Entwicklung von Computersystemen nach: Die Ära der anfänglichen Großrechner-systeme und Mainframes wurde bald durch die Vorherrschaft von Workstations oder PCs abgelöst – denen wiederum zurzeit der Rang als dominante Rechnerform von mobilen Systemen wie Notebooks oder PDAs abgelaufen zu werden scheint. Analog scheinen auch Projektoren den Sprung hin zum personalisierten Gerät zu schaffen – z.B. in Form von Heimkinoanlagen [2]. Auch der Trend zur Mobilität lässt sich deutlich erkennen: In Universitäten ist es inzwischen weitgehend üblich, dass Dozenten neben eigenen Notebooks auch eigene Projektoren mitbringen. Der tatsächliche Gebrauch des Projektors erfolgt aber weiterhin statisch.

Daneben finden sich aber auch Ansätze und erste Systeme, welche die Projektoren in weniger klassischer Weise verwenden:

- Automatisch kalibrierende Systeme: Von automatischer Keystone-Korrektur [6] über Anpassung an beliebige Projektionsflächen-Geometrien [1] bis hin zum Ausgleich optischer Eigenschaften der Projektionsfläche [13]. Erweiterung auf ad-hoc-Projektoren-Cluster und schwenkbare Projektoren [2].
- Systeme, welche die direkte Interaktion mit dem angezeigten Bild erlauben, z.B. Escritoire, SmartBoard oder Everywhere-Display (Näheres in Kapitel 2).
- Systeme, die Objekte in der Umgebung erkennen und auf sie reagieren [5][15].
- In der Hand gehaltene Systeme (siehe Kapitel 3).

...

Die folgende Arbeit gibt einen Überblick über Benutzerschnittstellen, die durch einen Projektor dynamisch auf verschiedenen Oberflächen erzeugt werden. Dabei liegt der Schwerpunkt auf Technologien, die keine weiteren technischen Maßnahmen außer dem Projektor (möglicherweise im Verbund mit einer Kamera) benötigen; insbesondere werden Systeme, die auf berührungsempfindlichen Oberflächen (TouchScreens) basieren ([17] gibt einen Überblick über diese) oder bei denen der Nutzer herkömmliche oder spezielle Eingabegeräte (wie Datenhandschuhe, Stifte, o.ä.) verwenden muss, nur am Rande betrachtet (Genaueres über diese Systeme sind in diversen anderen Arbeiten dieses Bandes zu finden). Im Wesentlichen werden zwei Geräteklassen vorgestellt: einerseits festinstallierte, z.T. schwenkbare Projektor-Kamera-Einheiten, welche eine oder mehrere Oberflächen des Raums, in dem sie installiert sind, als Projektionsfläche verwenden, um herkömmliche Bildschirmflächen zu ersetzen oder zusätzliche Informationen zu vorhandenen Objekten anzuzeigen; andererseits kleine, tragbare Projektoren im Verbund mit Kamera und ggf. weiteren Sensoren. Interaktion erfolgt bei letztgenannten durch Schwenken des Gerätes und Benutzung von auf diesem angebrachten Bedienelementen. Zu beiden Kategorien werden Prototypen, die wichtigsten Softwaretechniken (Bildkorrektur, Tracking, etc.) und Anwendungsbereiche vorgestellt, möglichst in der Reihenfolge aufeinander aufbauender Technologien.

Der rasante Fortschritt der angesprochenen Systeme und die Vielzahl der in diesem Bereich forschenden Teams lassen dabei nur eine unvollständige, exemplarischen Auswahl zu.

2 Statische Projektoren

Dieses Kapitel beschäftigt sich mit fest installierten Systemen, also Einheiten aus Projektoren und Kameras, die im laufenden Betrieb nicht bewegt werden. Dabei treten sehr unterschiedliche Konzepte und (mögliche) Anwendungsfelder zu Tage – ein deutliches Indiz für das Potential und die geringe Auslotung dieses Forschungsfeldes. Hier am Rande zu erwähnen sind Technologien, die zwar neu und innovativ sind, aber keine neuen Anwendungsbereiche eröffnen: Darunter fallen zum Beispiel aktuelle Software-Systeme, die es erlauben, das Bild mehrerer Projektoren (mit nicht zwingend gleicher Auflösung und Leuchtstärke) zu einem nahezu übergangslosen, größeren Gesamtbild zu vereinen – und das (nach ungefährender manueller Ausrichtung) mit Hilfe von Kameras fast automatisch [1]. Dennoch kann diese Technik nur eine altbekannte Aufgabe (nämlich die großflächige Darstellung von Bildern) besser erfüllen, aber per se keine neuen Verwendungszwecke eröffnen - Interessierten zu diesem Thema seien die Arbeiten von Aditi Majumder et al. empfohlen. Dasselbe gilt für die damit verwandte Technik „Virtuelle Rück-Projektion“ [2][10]: Mit Hilfe von mindestens zwei ausreichend weit voneinander platzierten Projektoren und einer Kamera wird ein Anstrahlen der Projektionsfläche simuliert, d.h. ein ausgeklügeltes Software-System kompensiert störende Schatten im Bild (siehe Abbildung 1), z.B. von einer vortragenden Person, durch dynamische Anpassung der Videosignale zu den Projektoren. Dennoch werden die Geräte weiterhin „nur“ zur visuellen Unterstützung von Präsentationen gebraucht.



Abbildung 1: Schattenkompensation durch virtuelle Rückprojektion [10]

2.1 Tele-Graffiti

Das „Tele-Graffiti“ [3] getaufte System verwendet ähnliche Technologien, setzt diese aber für einen anderen Zweck ein. Der Prototyp besteht aus einem Projektor, der mit Hilfe eines Spiegels Bildsignale auf einen Zeichenblock werfen kann, und einer Kamera, die auf den selben Block ausgerichtet ist. Diese Konstruktion ist mit einem PC verbunden, der mit einem entfernten PC, an den der selbe Aufbau angeschlossen ist, per Netzwerk kommuniziert (siehe Abbildung 2).

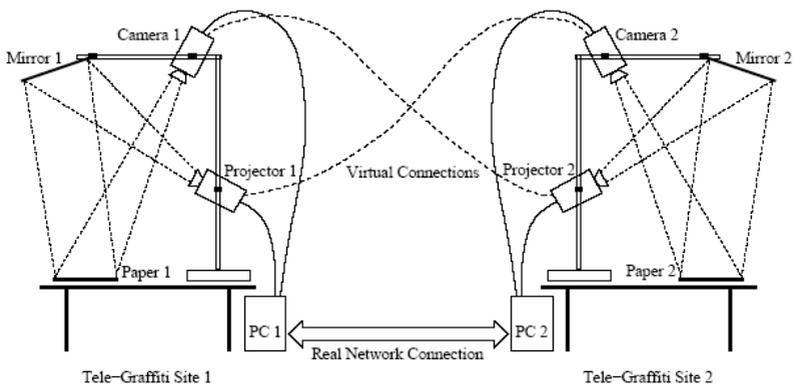


Abbildung 2: Schematischer Aufbau des Tele-Graffiti-Systems [3]

Die Idee dahinter ist nun, dass zwei räumlich von einander getrennte Benutzer auf natürliche Weise an der selben Zeichnung arbeiten können: Das auf einer Seite aufgenommene Bild wird über das Netzwerk übertragen und auf den Zeichenblock des Gegenübers projiziert (und umgekehrt). So soll auf beiden Zeichnungen ungefähr das gleiche Bild zu sehen sein.

Eine besondere Herausforderung bei diesem und vielen anderen Systemen, bei denen eine Kamera eine projizierte Fläche aufnimmt, stellt dabei das Problem der Rückkopplung [3] dar: Die Kamera nimmt sowohl tatsächliche Zeichnungen, als auch das Bild von einem Projektor und einem vielleicht gerade zeichnenden Benutzer, auf und schickt es zur Anzeige an den Projektor der Gegenseite – wo es wieder aufgenommen wird, usw. Damit werden „Altlasten“ wie z.B. die Hand des Benutzers immer wieder hin und her übertragen; das angezeigte Bild wird so schnell unbrauchbar. Selbst ohne menschliche Eingriffe ist so eine Rückkopplungskatastrophe möglich: Ist die Helligkeitswiedergabe der Kameras nicht exakt eingestellt, wird das Bild schnell heller oder dunkler bis das Auflösungsvermögen der Kamera überschritten ist – somit ist in beiden Fällen das System nicht mehr verwendbar, wenn auch der zweite Effekt durch Software-Eingriffe nahezu ausgeschlossen werden kann: In [3] wird ein Algorithmus vorgeschlagen, der einen erwarteten Helligkeitswert für jeden Pixel berechnet und die übertragenen Bilder entsprechend anpasst.

Dennoch bleibt das System höchst instabil und ist in der Praxis kaum verwendbar: Jede Helligkeitsänderung im Zeichenbereich (wie eine an- oder abgeschaltete Schreibtischlampe, ein Nutzer, der sich über das Blatt beugt, ...) kann zu einer Rückkopplungskatastrophe führen. Eine mögliche Abhilfe wäre die Verwendung von Stiften mit spezieller fluoreszierender Tinte, sodass eine mit entsprechendem Filter ausgestattete Kamera nicht von der Projektion beeinflusst wird; die Möglichkeit der Farbwiedergabe wird dadurch aber genommen. Bei Verwendung einer anderen Technologie zur Erkennung der Zeichnungen, wie z.B. berührungsempfindliche Oberflächen, besteht die Gefahr, dass der große Vorteil für den Nutzer, nämlich auf natürliche Weise zeichnen zu können, verloren geht.

Darüber hinaus ist noch zu beanstanden, dass keinerlei Interaktionsmöglichkeiten mit dem System vorgesehen sind, herkömmliche Ein- und Ausgabegeräte sind also nach wie vor notwendig – ein Problem, welchem der folgende Entwurf begegnet.

2.2 Einzelplatz-Projektor

Das sogenannte „Escritoire“¹-System [4] ist ein Vertreter der „personal projected displays“ (kurz: PPD, im Deutschen ungefähr „Einzelplatz-Projektor“), also ein Digital-Projektor für einen Arbeitsplatz.

Der Prototyp besteht aus zwei Standard-Projektoren, die wie in Abbildung 3 angeordnet sind. Dabei bestrahlt der unter dem Schreibtisch aufgestellte DP über einen Spiegel die gesamte Schreibtischfläche, während der zweite DP oberhalb des Schreibtischs nur auf einen kleinen Bereich auf der Arbeitsfläche fokussiert. Ein PC mit moderner 3D-Grafikkarte und zwei Monitoranschlüssen speist hierbei beide Projektoren. Für die Interaktion stehen ein per Ultraschall georteter Stift und ein großer Digitalisierer (ähnlich einem Grafik-Tablett) zur Verfügung.

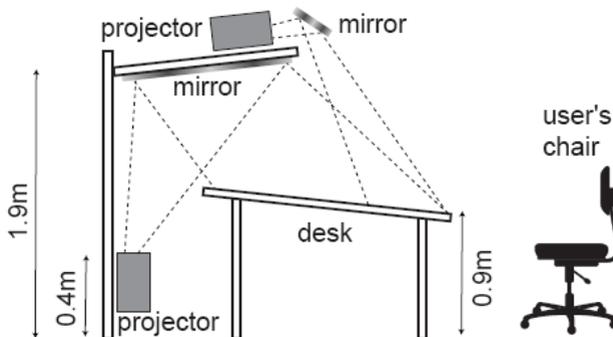


Abbildung 3: Schematischer Aufbau des "Escritoire"-Systems [4]

Eine Besonderheit bei dem „Escritoire“-System ist, dass weder Spezialprojektoren noch besondere Optiken zur Entzerrung des Bildes verwendet werden; diese Korrekturen werden durch Software und die 3D-Karte vorgenommen: Ist eine Rekalibrierung der Anzeige notwendig, werfen beide Projektoren nacheinander mehrere vom Benutzer zu berührende Ziele auf die Arbeitsfläche [4]. Mit Hilfe des Digitalisierers kann so die normale Verzerrung bestimmt und durch entsprechende Anweisungen an die Grafikkarte ausgeglichen werden. Anders ausgedrückt werden Daten über die Umgebung gesammelt und für die Bildwiedergabe verwendet – ein Konzept, das in später vorgestellten Systemen wieder auftauchen wird.

Der Ansatz mit zwei Projektoren und entsprechenden Programmen ermöglicht dem Nutzer, die Arbeitsfläche ähnlich wie einen normalen Schreibtisch zu verwenden: Mit

¹ Französisch: Schreibtisch

Hilfe des Ultraschall-Stifts können beliebige Dokumente wie ein Blatt Papier auf der Projektionsfläche des PPD umhergeschoben und sogar gestapelt werden. Möchte der Anwender ein Dokument genauer betrachten oder bearbeiten, schiebt er es von der grob aufgelösten Peripherie (also der Fläche, die nur von dem DP hinter dem Tisch beschienen wird) in das durch den zweiten Projektor höher aufgelöste Zentrum. Hier können nun mit dem von dem Digitalisierer georteten Stift Eingaben oder Anmerkungen gemacht werden.

Eine Anwendung für das System ist die gemeinsame Arbeit an einer Zeichnung wie im vorher vorgestellten „Tele-Graffiti“ durch mehrere verbundene PPDs – ohne dass die vorher beschriebenen Rückkopplungseffekte auftreten können; das Zeichnen erfolgte bei Letztgenanntem allerdings natürlicher. Außerdem macht die beschriebene Arbeitsweise das Escrioire besonders für Personen interessant, deren Hauptaufgabe in der Auswahl und Kommentierung von Dokumenten besteht (z.B. Lektoren oder Chefredakteuren). Durch entsprechend angepasste Programme kann das Anwendungsfeld auch auf Grafiker ausgeweitet werden, die häufig bereits mit Grafik-Tabletts arbeiten und sich so schnell an den ähnlichen Digitalisierer gewöhnen dürfen.

Für einen tatsächlichen Praxis-Einsatz spricht dabei der geringe technische und finanzielle Aufwand: Digitale Projektoren und die angesprochenen Ortungssysteme sinken beständig im Preis, die Anforderungen an den PC liegen unterhalb der zur Zeit verfügbarer Standardhardware und der Restaufbau (Tisch, Spiegel, Projektorhaltungen) muss neben einer gewissen Stabilität keine besonderen Eigenschaften aufweisen, sodass der Gesamtpreis für die Apparatur unter 5.000 € liegen dürfte (vergleiche [4]). Gegen eine allgemeine Verbreitung spricht der hohe Spezialisierungsgrad: Eine der am häufigsten genutzten Computer-Anwendungen ist die Eingabe von Text, und genau hierfür scheint das Escrioire-System aufgrund noch zu langsamer und fehleranfälliger Programme zur Handschriftenerkennung ungeeignet. Allerdings sind auch Lösungsansätze wie eine ausziehbar unter dem Tisch angebrachte Tastatur denkbar. Der Schattenwurf der eigenen Hände wurde von Testpersonen bei einer durchgeführten Evaluation [4] als kaum störend empfunden und bedarf somit auch keiner Lösungsstrategien. Bisher ungelöste Kritikpunkte sind die starke Wärmeabgabe und Lärmbelästigung durch die Projektoren – ein deutlicher Störfaktor für konzentriertes Arbeiten.

2.3 SearchLight

Alle bisher vorgestellten Ansätze und Systeme haben als Gemeinsamkeit, dass die Projektions- und Eingabefläche während des Gebrauchs unverändert an der selben Stelle verharrt und somit auch den Nutzer an einen Ort bindet. Zur Einführung flexiblerer Konzepte wird nun das „SearchLight“ [5] (dt.: „Suchscheinwerfer“) vorgestellt.



Abbildung 4: Prototyp des "SearchLight"-Systems aus [5]

Der Prototyp besteht aus einer Einheit aus einem handelsüblichen DP und einer daran befestigten Kamera, die zentral an der Decke in einem Raum fest installiert ist. Diese Einheit ist so aufgehängt, dass sie um die räumliche Hochwertachse rotieren und nach oben und unten geschwenkt werden kann (siehe Abbildung 4). Dadurch ist nahezu jeder Ort im Raum – sofern keine Gegenstände wie z.B. Regale im Weg stehen – von der Kamera aufnehmbar und vom Projektor beleuchtbar.

Durch optische Marker (und adäquate Software) wird dieser Aufbau nun zu einer Suchfunktion für reale Objekte aufgewertet: Jedes Objekt, das von dem System in Zukunft gefunden werden soll, erhält einen individuellen optischen Marker an seiner der Raummitte zugewandten Seite. Die Zuordnung zwischen Objekt und Marker wird zusätzlich in einer Datenbank festgehalten. Die Software, die den „Suchscheinwerfer“ ansteuert, lässt diesen nun zeilenweise den ganzen Raum „absuchen“ und nimmt eine Folge überlappender, hochauflösender Fotos auf. Diese Bilder, zu denen die korrelierenden Rotations- und Neigungswinkel gespeichert werden, durchsucht das Programm anschließend nach den vergebenen optischen Markern und speichert die Position des Markers im Bild zusammen mit den Rotations- und Neigungsdaten zu diesem Bild in der Datenbank der suchbaren Objekte. Die Fotografie kann nun verworfen werden. Startet ein Anwender nun eine Suchanfrage nach einem gefundenen Objekt (bei nicht gefundenen erhält er eine Fehlermeldung), stellt die Software die Projektor-Kamera-Einheit auf die gespeicherten Rotations- und Neigungswerte ein und lässt den Projektor dann an der gespeicherten Position im aufgenommenen Bild eine optische Hervorhebung anzeigen (z.B. eine leuchtende Ellipse) [5]. Wurde das Objekt seit dem letzten Suchvorgang nicht bewegt, sollte es nun beleuchtet und damit schnell auffindbar sein.

Ein tatsächlicher Einsatz des Systems in realen Umgebungen scheint unwahrscheinlich, da nur sehr spezielle Umgebungen und Objekte in Frage kommen: Die markierten Objekte müssen eine bestimmte Größe haben, um überhaupt markiert werden zu können, und sie müssen immer in der selben Orientierung (nämlich mit dem Marker zur Raummitte) im Raum platziert werden; Bibliotheken würden diese Anforderungen erfüllen. Allerdings sind Bibliotheken normalerweise mit stehenden Regalen ausgestattet, wodurch manche Bücher einfach ignoriert oder für jede Regal-

reihe „SearchLights“ installierten werden müssten. Außerdem ist die Anzahl zuverlässig unterscheidbarer optischer Marker durch deren geringe Größe und perspektivische Verzerrungen der Fotos begrenzt und wahrscheinlich für Bibliotheken nicht ausreichend. Darüber hinaus ist das System durch unachtsame Nutzer, die die Gegenstände falsch orientiert oder verdeckt ablegen, verwundbar. Zuletzt bedingt die Pflege der Objekt-Datenbank einen deutlichen Wartungsaufwand. Positiv herzuheben ist allerdings, dass das System sonst keinerlei Kalibrierungsarbeiten oder dreidimensionaler Nachmodellierung des Raums notwendig macht.

2.4 Tragbare Projektionsflächen

Das nun vorgestellte System der „tragbaren Projektionsflächen“ [6] gleicht im technischen Aufbau dem gerade beschriebenen „Suchscheinwerfer“ (siehe Abbildung 4), erreicht aber durch andere Programme und Algorithmen ein neues Anwendungsgebiet und lässt auch direktere Interaktion mit der Projektion zu. Allerdings erfordert es für die anfängliche Kalibrierung weitere Kameras.

Für das Problem der perspektivischen Korrektur von projizierten Bildern werden bei diesem System gleich zwei Lösungsstrategien verwendet. Vorauszuschicken ist, dass gerade Linien, die von einer Lichtquelle aus auf eine gerade Oberfläche perspektivisch projiziert werden, wieder auf Geraden abgebildet werden (von einem beliebigen Standort aus gesehen). Die Transformation, die die Gerade dabei erfährt, kann durch eine invertierbare 3×3 -Matrix dargestellt werden. Durchläuft die Gerade vor der Projektion die zu dieser inverse Matrix, wird die Gerade so dargestellt, als würde die Oberfläche senkrecht bestrahlt [6].

Mit diesem Vorwissen und der Erkenntnis, dass die meisten Oberflächen in Räumen eben sind, haben S. Borkowski et al. [6] folgendes Verfahren in zwei Phasen zur Darstellung entzerrter Bilder auf beliebigen (ebenen) Oberflächen in einem Raum entwickelt: Die Projektor-Kamera-Einheit fährt wie beim „SearchLight“ den Raum in diskreten Schritten zeilenweise ab und strahlt dabei ein Gittermuster aus farblich codierten Linien aus. Weitere im Raum befindliche Kameras (für eine Vollautomatisierung des Kalibrierungsprozesses sind weitere steuerbare Kameras nötig!) nehmen die bestrahlten Flächen auf und geben die Bilder an den zentralen Rechner weiter. Dort erkennen spezielle Algorithmen, wo die Linien „abknicken“; aus der farblichen Kodierung der Linien kann nun bestimmt werden, wo die Umrisse einer Oberfläche liegen. Diese Informationen werden zusammen mit Rotations- und Neigungsdaten des Projektors gespeichert und bilden so eine genaue Umgebungskarte von Oberflächen unterschiedlicher Orientierung. Zur Feststellung, welche genaue Orientierung eine Oberfläche hat, wird ein Algorithmus verwendet, dessen Erklärung den Rahmen dieser Arbeit sprengen würde – genauere Informationen finden sich unter [6], [18] und [19]. Dieser Algorithmus verwendet allerdings auch nur die angeführten Gerätschaften und liefert als Ergebnis hinreichend genaue Transformationsmatrizen für jede gefundene Oberfläche ausreichender Größe.

Das gerade vorgestellte Verfahren ist automatisierbar und sehr flexibel, aber prinzipiell langsam und ungeeignet, perspektivische Korrekturen für mobile Projektionsflächen zu errechnen. Für diese Aufgabe verwenden S. Borkowski (et al.) eine andere Lösung. Diese braucht zwar keine weiteren Kameras im Raum außer der auf dem DP

montierten, stellt dafür aber Ansprüche an die verwendete Projektionsfläche: nur rechteckige, helle Flächen mit deutlich abgesetztem dunkleren Rand (siehe Abbildung 5) können erkannt werden. Wie im ersten Verfahren ist auch hier das Ziel, eine Transformationsmatrix für die Projektionsfläche anzugeben. Durch die Einschränkungen bei den geeigneten Zielen der Projektion reicht es aber nun, die vier Ecken der Fläche zu verfolgen, da aus diesen die gesuchte Matrix direkt erstellt werden kann. Sind die Ecken erkannt, übernimmt ab dann ein optischer Algorithmus (eine genaue Beschreibung findet sich unter [6]) die Verfolgung derselben. Zusätzlich wird der Schwerpunkt dieser vier Punkte berechnet und seine Abweichung vom Zentrum des aufgenommenen Bildes zur Nachführung der Projektor-Kamera-Einheit herangezogen. Somit wird das darzustellende Bild beständig an die herumgetragene Fläche angepasst.



Abbildung 5: *links:* eine tragbare Projektionsfläche; *rechts:* die selbe Fläche im mobilen Betrieb [6]

Das Zusammenspiel dieser beiden Techniken kann nun für den virtuellen Transport des darzustellenden Bildes eingesetzt werden: Angenommen, eine beliebige statische Oberfläche im Raum wird als Bildschirm verwendet, der Nutzer möchte aber, dass das Bild auf einer anderen Oberfläche dargestellt wird, dann genügt es, eine portable Fläche, die den oben angegebenen Restriktionen entspricht, in den Erfassungsbereich der Kamera einzuführen. Das System erkennt die portable Fläche und „überträgt“ die Projektion auf diese (siehe Abbildung 5, rechts). Nun trägt der Anwender den tragbaren Bildschirm zur gewünschten Ziel-Oberfläche und verharrt dort für einen Augenblick. Dies wertet das System als weiteren Übertragungswunsch und stellt von nun an das Bild auf der Fläche dar, die der größte Teil des getragenen Schirms überdeckt.

Eine Verwendung des Systems für sich genommen ist in der Praxis noch kaum zu erwarten: Zwar sind die Anschaffungskosten der nötigen Geräte gering, v.a. da nur Standard-Produkte benutzt werden, und die vorgestellten Technologien funktionieren nachweislich – eine tatsächliche Anwendung ist aber nicht auszumachen. Der Einsatz als Substitution für herkömmliche Monitore ist zwar denkbar, dürfte sich aber in der Praxis kaum lohnen.

Vor allem bestimmte Schwächen des Systems erschweren einen sinnvollen Einsatz. So ist zwar die Darstellung auf nahezu allen Oberflächen des Raums möglich,

aber nur für wenige sinnvoll: Die perspektivische Korrektur wird durch eine Vorverzerrung des Bildes erreicht, d.h. nur noch ein kleinerer Teil des Auflösungsvermögens des Projektors kann für die Darstellung verwendet werden – die tatsächliche Bildauflösung sinkt. Zudem bedingt die Vorverzerrung bei manchen vormals geraden Linien ein Kippen in eine bestimmte Richtung, was sich vor allem bei Text als störend erweist. Diese Effekte verschlimmern sich je spitzer der Winkel zwischen Oberfläche und optischer Achse des DPs wird. Die verbliebene Zahl adäquater Projektionsflächen wird weiterhin durch die optischen Eigenschaften der Oberfläche reduziert: Wild gemusterte Teppiche sind so ungeeignet wie (saubere) Glastische oder Wandpaneele in Klavierlack-Optik. Da außerdem die Interaktion mit dem System (außer dem „Übertragen des Bildes“) weiterhin über klassische Eingabemedien erfolgt, ist der Nutzer für klassische Interaktionen an seinen normalen Arbeitsplatz gebunden. Besonders hervorzuheben ist dennoch die einfache Kalibrierung im Vergleich mit anderen Systemen, die eine vollständige dreidimensionale Nachmodellierung des Raums benutzen und dementsprechend unflexibler sind, und der intuitive Wechsel zwischen den Projektionsflächen. Wie viel Potenzial der Ansatz tatsächlich hat, zeigt der folgende Prototyp.

2.5 Der „Everywhere Displays Projektor“

Der „Everywhere Displays Projektor“ (dt.: „Überall Bildschirme Projektor“, im Folgenden mit „EDP“ abgekürzt) wurde von IBM als Teil des Gesamtkonzeptes „BlueSpace“ [7], einem Entwurf des Büros der Zukunft, vorgestellt und seitdem fortlaufend weiterentwickelt [8][19]. Tatsächlich ist der EDP das erste Konzept gewesen, bei dem die Projektionsfläche nicht statisch ist, sondern während des Betriebs dynamisch und automatisch geändert wird – so ist durch den EDP das Forschungsfeld der steuerbaren Projektoren überhaupt erst eröffnet und die Grundlage für die zuletzt vorgestellten Prototypen geschaffen worden. Im Folgenden wird nur auf den zuletzt präsentierten Entwicklungsstand des Systems eingegangen. Der Aufbau der Projektor-Kamera-Einheit folgt demselben Ziel wie bei dem vorher beschriebenen Entwurf, nämlich der Projektion von Computerausgaben auf möglichst alle Flächen eines Raumes, erreicht dies aber über einen dreh- und schwenkbaren Spiegel (siehe Abbildung 6). Angesteuert wird der EDP von einem zentralen Computer, der auch die anderen Bestandteile des „BlueSpace“-Konzeptes verwaltet [7].

Die Bildkorrektur wird durch eine Standard-3D-Grafikkarte übernommen, die auf ein dreidimensionales Modell des Raums zurückgreift. Hierbei wird der Umriss des darzustellenden Bildes als Textur auf die virtuelle Entsprechung der realen Projektionsfläche gelegt und von einer virtuellen Kamera, die sich möglichst genau an der Position des realen EDP befindet aufgenommen. So erhält man einen perspektivisch verzerrten Umriss des Bildes, der wie im Abschnitt 2.3 zur Erstellung einer Transformationsmatrix des Bildes genutzt wird [7][19].

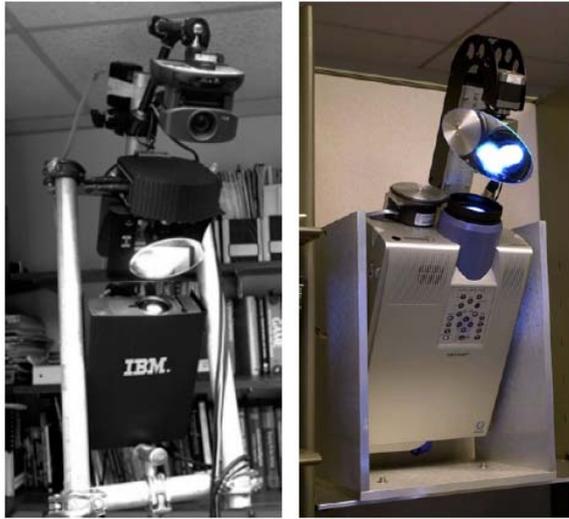


Abbildung 6: Zwei Prototypen des "Everywhere Displays Projector" [8]

Hinzu kommt, dass das EDP-System Handerkennungs-/verfolgungs- Algorithmen mit einbezieht, wodurch eine direkte Interaktion mit dem projizierten Bild möglich ist [8]. Diese Algorithmen arbeiten rein optisch, indem sie abweichende Helligkeiten in einzelnen aufgenommenen Bildern und zwischen aufeinanderfolgenden Bildern erkennen und zu einer Folge von schwarz/weiß- Bildern (schwarzer Pixel: keine Auffälligkeiten, weißer Pixel: hier ist eine Abweichung oder Änderung erkannt worden) erzeugt. Diese Daten werden nun aufbereitet und das Ergebnis mit einer Finger- oder Handvorlage verglichen (siehe Abbildung 7. Eine detailliertere Einführung gibt [11]) [8]. Auf diese Weise kann nicht nur die Hand eines Benutzers verfolgt, sondern auch Gesten erkannt werden: Ein plötzliches (optisches) Verkürzen gefolgt von einem Strecken des Zeigefingers, könnte so als Betätigen einer Schaltfläche interpretiert werden.



Abbildung 7: Ablauf einer Finger-Erkennung [8].
 Links: aufgenommenes Bild. Mitte: Änderungsanalyse.
 Rechts: Interpretation der Änderung als Finger.

Ähnliche Algorithmen werden im Rahmen des EDP-Systems auch verwendet, um herauszufinden, wo sich der oder die Nutzer gerade befinden, allerdings mit größeren

Toleranzrahmen und ungenauem Ergebnis: Die Erkennung der exakten Haltung und Orientierung eines Menschen im Raum anhand von zweidimensionalen Aufnahmen überfordert immer noch jeden bekannten Bilderkennungsalgorithmus [8].

Für das EDP-System gelten die selben Einschränkungen für verwendbare Projektionsflächen, die im Abschnitt 2.3 vorgestellt wurden. Hinzu kommt noch die weitere Restriktion, dass bei manchen Flächen der Nutzer zwischen Projektor und Bild steht, und somit effektives Arbeiten durch seinen eigenen Schattenwurf beeinträchtigt. Außerdem ist der EDP-Ansatz zwar relativ flexibel, sobald die aufwändige Einrichtung erfolgt ist, kann aber nicht mit Änderungen der Inneneinrichtung des Raumes umgehen: Werden die Möbel umgestellt, ist eine erneute Modellierung notwendig (Alternativen zu dieser Modellierung werden aber bereits erforscht und getestet [19]). Zudem wird damit von den Anwendern erwartet, dass sie ihre Räumlichkeiten in penibler Ordnung halten, da herumliegende Gegenstände und Papiere vom System nicht erkannt werden und so die Bildwiedergabe beeinträchtigen. Daneben können die Gestenerkennungsalgorithmen nur vernünftige Ergebnisse liefern, wenn die Hand sehr nah an der Projektionsfläche ist und/oder die Fläche fast senkrecht bestrahlt wird – ansonsten ist Erkennung des Ortes von Gesten (z.B. welchen Button der User gerade drückt) durch fehlende Tiefeninformationen sehr ungenau.

Trotz dieser Schwächen, eröffnet die EDP-Technologie eine Vielzahl von Anwendungsfeldern. Im Rahmen eines Büros der Zukunft kann der steuerbare Projektor als Großbildschirm für spontane Besprechungen mit Mitarbeitern verwendet werden. Oder er ergänzt den bestehenden Monitor als zweiter, berührungsempfindlicher Bildschirm. Wird das Gerät gerade nicht verwendet, ist eine Anwendung als Gestaltungsmittel denkbar, wie z.B. der Anzeige eines virtuellen Fensters (vgl. Abbildung 8). Außerdem ist das System dazu in der Lage, Informationen in einer unaufdringlichen Art im peripheren Blickfeld des Anwenders zu positionieren [9]. Die Technologie kann aber auch außerhalb von Büros sinnvoll eingesetzt werden, z.B. zur Unterhaltung wie die „Fro..og!“-[19]-Anwendung: Der Benutzer soll einen Frosch, der auf einer beliebigen Fläche im Raum angezeigt wird, fangen, indem er auf ihn tippt (vgl. Abbildung 8). Das System erkennt die entsprechende Handbewegung und lässt den Frosch zu irgend einer anderen Oberfläche im Raum „springen“. Das Programm ist zwar bisher rudimentär, zeigt aber die Möglichkeiten auf. Ähnlich geartete Anwendungen bieten sich als neue, intuitivere Lernsysteme für Kinder an. Im industriellen Bereich könnten EDP eingesetzt werden, um die nächsten Arbeitsschritte zu verdeutlichen, usw.



Abbildung 8: Anwendungen des ED-Projektors. Links: Darstellung eines virtuellen Fensters [19]. Mitte, rechts: Ein Benutzer spielt „Fro..og!“ [19]

2.6 Weitere Technologien

Abschließend für dieses Kapitel sollen nun noch zwei vielversprechende Technologien vorgestellt werden, die die Möglichkeit bieten, die Nutzbarkeit der bisher vorgestellten Prototypen weiter zu verbessern.

Die meisten bekannten Systeme zur Gestenerkennung müssen vor deren ersten Gebrauch von Service-Personal angelernt werden. Dabei stellen sich die Systeme tatsächlich auf die anlernende Person ein und zeigen mit dieser die größte Zuverlässigkeit bei der Erkennung. Ein neuer Ansatz von A. Licsár (et al.) [12] erlaubt nun das dynamische Training des Systems während der Anwendung, sodass auch bei ungeübten und häufig wechselnden Nutzern befriedigende Ergebnisse erzielt werden können – und das bei neun verschiedenen Gesten, die aus einer Kombination von Unterarm-, Handflächen- und Fingerbewegungen bestehen. Die dynamisch Anpassung an den Nutzer beruht auf der Einführung einer zusätzlichen Geste, die dem System mitteilt, dass eine Bewegung falsch erkannt wurde. Die Software wechselt dann in den Trainings-Modus, in dem der Nutzer die fragliche Geste wiederholen soll, um dem Programm die neuen Parameter für diese einzugeben. Wurde eine Geste richtig erkannt, d.h. die „Falsch-Erkant“-Geste wird nicht eingegeben, übernimmt das System die Parameter von der zuletzt eingegebenen Bewegung[12].

Eine weitere neue Technologie soll Projektionen unabhängiger von der verwendeten Projektionsfläche (und damit fast universal einsetzbar) machen: Anders als frühere Ansätze, die störenden Oberflächenbeschaffenheiten durch vorher festgelegte Modelle der Fläche entgegen wirken, nähert sich der vorgeschlagene Algorithmus von S. Nayar (et al.) [13] in mehreren Durchläufen einem ausgeglichenen Bild an, indem die Helligkeit (bei farbigen Bildern die Intensität einer Farbe) jedes Pixels betrachtet und mit der durchschnittlichen Helligkeit des erwünschten Bildes, des projizierten Bildes und der Helligkeit des Pixels im Originalbild verrechnet wird. Daraus ergeben sich grobe kompensierte Pixelhelligkeiten, die nun angezeigt werden. Der Vorgang beginnt von vorne, bis eine hinreichende Übereinstimmung erreicht ist [13].

3 Mobile Projektoren

Das vorausgegangen Kapitel war ausschließlich auf statische Systeme beschränkt, die z.T. durch Motoren bewegt, Bilder auf verschiedene, ebenfalls meist statische Oberflächen projiziert haben. Die Prototypen dieses Kapitels gehen einen Schritt weiter: Sie sind so konzipiert, dass sie von Benutzern einfach mitgenommen und im mobilen Betrieb verwendet werden können.

3.1 Gemischte Realität mit einem Rucksack-Projektor

Der erste hier vorgestellte Prototyp bringt alle Systembestandteile (DP, Kamera und Beleuchtungseinheit mit Infrarotfilter und mobilem PC) in einem Rucksack unter, sodass Projektor und Kamera über die Schulter des Benutzers hinweg auf eine Fläche vor diesem fokussieren (siehe Abbildung 9) [14].



Abbildung 9: Der Rucksack-Projektor in Aktion [14]

Zur Darstellung korrekter Bilder ist das System auf optische Marker, die an der Projektionsfläche angebracht sind, angewiesen. Dazu werden spezielle, das von der Lampeneinheit ausgestrahlte Infrarotlicht reflektierende Markierungen verwendet. Diese sind an den vier Ecken der notwendig rechteckigen Projektionsfläche angebracht und definieren so den Umriss des darzustellenden Bildes. Der Algorithmus zur Korrektur perspektivischer Verzerrungen arbeitet analog zu dem Algorithmus zur Erkennung tragbarer Projektionsflächen aus Abschnitt 2.4. Zusätzliche in bestimmten Mustern angebrachte Markierungen dienen als Identifikatoren und weisen so einer Oberfläche ein Anwendung zu. Interaktionsmöglichkeiten werden durch eine am Zeigefinger angebrachte Infrarot-Leuchtdiode eröffnet – inwieweit damit Gesten wie das Antippen einer Schaltfläche erkannt werden können, konnte bis jetzt nicht evaluiert werden [14].

Der bestehende Prototyp beinhaltet bisher zwei praktische Anwendungen: Einerseits ein Ringbuch, dessen einzelne Seiten mit verschiedenen Identifikatoren für unterschiedliche Applikationen wie Videobetrachtung oder ein Kartenspiel versehen sind. Andererseits ein personalisiertes schwarzes Brett: Das System erkennt an den optischen Markern des „schwarzen Bretts“ (eine ansonsten leere Fläche) die verlangte Applikation sowie Umgebungsvariablen und projiziert auf den Nutzer abgestimmte Informationen auf die Fläche [14].

Der prinzipielle Ansatz des Systems, vor dem Nutzer befindliche Gegenstände mit zusätzlichen Inhalten zu erweitern, klingt vielversprechend, ist aber bisher nur rudimentär umgesetzt: Die Verwendung von groben Mustern als Identifikatoren nimmt noch zuviel Platz in Anspruch und scheint kaum skalierbar. Ebenso genügt die Interaktion mit dem System kaum den Anforderungen eines praktischen Einsatzes. Erschwerend kommt noch die Unhandlichkeit und das Gewicht der Apparatur hinzu. Allerdings scheinen weiterentwickelte Systeme dieser Art gegenüber anderen Versuchen, die Realität durch computergenerierte Inhalte zu erweitern (wie z.B. Datenhelmen mit halbdurchlässigen Bildschirmen) zwei Vorteile zu besitzen: Der Nutzer kann direkt mit der projizierten Schnittstelle durch Gesten interagieren und anderen Personen seine gespeicherten Daten zeigen.

3.2 Der objekt-adaptive Bildschirm

Einen ähnlichen Ansatz stellt das „object-adaptive Display“ [2] (dt.: objekt-adaptiver Bildschirm, kurz: OAD) dar: Der OAD verwendet ebenfalls optische Marker zur Identifikation von Objekten und zur perspektivisch korrekten, ortsangepassten Darstellung von Informationen. Dafür werden Marker eingesetzt, die an Tortendiagramme erinnern und sich untereinander durch Farbgebung und Größe der einzelnen Segmente unterscheiden; so sind mehrere tausend von diesen eindeutig identifizierbar. Der OAD-Prototyp besteht im Wesentlichen aus einer Einheit aus Projektor, Kamera, Computer und Neigungssensor, und ist dazu gedacht, in der Hand gehalten und getragen zu werden. Dies ermöglicht die Integration einer mausartigen Bedienung: Weil das projizierte Bild durch die Marker an eine bestimmte Stelle gebunden ist, bewirkt ein Schwenken des Gerätes nur eine Verschiebung des Bildes im Pixel-Raster des Projektors – mit anderen Worten: Der Prototyp kann in bestimmten Grenzen gedreht und geneigt werden ohne das Bild zu beeinflussen. Lässt man nun einen Maus-Zeiger an einer festen Position auf der Bildfläche des Projektors darstellen, kann man durch Drehen und Schwenken des Gerätes einen Zeiger über die angezeigt Benutzerschnittstelle bewegen; bringt man nun noch Tasten an dem OAD an, stehen sämtliche Funktionen zur Verfügung, die auch mit einer Maus erledigt werden können. Dabei scheint das Zeigen mit dem Gerät intuitiver und natürlicher als indirekte Eingaben mit einer klassischen Maus – wenn auch die Bedienungs-Geschwindigkeit hinter dieser zurückstehen dürfte[2].

Der OAD, bzw. ein ausgereifterer Nachfolger, ist auf dem besten Weg, praktischen Einsatz zu finden. Das System dient in erster Linie dazu, zusätzliche Informationen zu erkannten Objekten anzuzeigen – diese Eigenschaft könnte in Verbindung mit einem stetigen Abruf neuer Daten (wie von Raskar et al. Vorgeschlagen [2]) überaus nützlich zur Verwaltung größerer Lagerhallen oder visueller Markierung von Rohr- oder Kabelverläufen bei Umbauarbeiten in Gebäuden sein. Darüber hinaus kann der OAD auch für die Bedienung einfacher Nutzerschnittstellen benutzt werden, wodurch die Verwendung als PDA-Ersatz denkbar wäre – aber erst müssen OAD-artige System deutlich leichter und handlicher werden.

3.2 RFIG Lamps

Der „RFIG Lamps“ [15] getaufte Prototyp ähnelt im generellen Aufbau und der Funktionsweise dem OAD, besitzt aber zusätzlich eine Empfangseinheit für Radiowellen, einen Beschleunigungssensor und ein Gyroskop (siehe Abbildung 10).

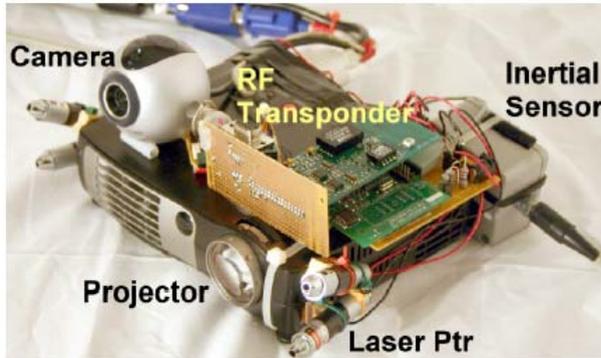


Abbildung 10: Das Innenleben des "RFIG-Lamps"-Systems [15]

Statt optischen Markern verwendet das System aber „Radio Frequency Identity and Geometry“-Marker. Dabei handelt es sich um intelligente, mit einem Photosensor ausgestattete Mikro-Computer, die nach Aktivierung über ein Radio-Signal Ihre Identifikationsnummer und die im Moment erhaltene Lichtintensität über Radiowellen aussenden; die Intensität bei der Aktivierung wird dabei als 0 bestimmt. Diese Eigenschaft ermöglicht es dem Prototypen, die räumliche Lage der einzelnen Marker festzustellen, indem er eine Folge binärer Lichtmuster sendet, sodass jeder bestrahlte Pixel von einer eindeutigen Folge von Helligkeitswerten beleuchtet wird. Der Hauptvorteil dieser aufwändigeren Technik gegenüber optischen Markern ist die nahezu unbegrenzte Verfügbarkeit an Markern mit unterschiedlichen Identifizierungsnummern [15].

Eine andere Weiterentwicklung gegenüber dem OAD ist die „quasi-stabilisierte Projektion“, d.h. mit Hilfe der von Beschleunigungssensor und Gyroskop gelieferten Daten können Schwenks des Gerätes auch ohne optische Bezugspunkte erkannt werden. Dadurch wird es möglich, die Projektion auf einer ebenen Fläche zu bewegen, ohne dass perspektivische Verzerrungen auftreten. Alternativ kann mit dieser Technologie die Projektion auch ohne Bezugspunkte nahezu stabil gehalten werden, was insbesondere die Verwendung des Gerätes auf jeder beliebigen (ebenen) Oberfläche gestattet [15].

Die Anwendbarkeit und der Nutzen der Technik der tragbaren Projektoren in der Praxis ist mit diesen neuen Technologien weiter verbessert worden, das Problem des hohen Gewichts und der Unhandlichkeit ist allerdings geblieben. Zudem ist der Prototyp nach wie vor an ein Kabel zur Stromversorgung gebunden. Diese Kritikpunkte konnten bei folgendem Entwurf nahezu beseitigt werden.

3.4 Spotlight Navigation

Der „Spotlight Navigation“ genannte Prototyp ist in einem kleinen und leichten Gehäuse untergebracht (siehe Abbildung 11). Technische Details sind nicht verfügbar, aber es ist anhand der Funktionalität des Geräts davon auszugehen, dass neben dem offensichtlichen, monochromen Projektor und obligatorischer Computer-

Hardware noch ein Beschleunigungsmesser und ein Gyroskop verwendet werden. Allerdings besitzt das System weder eine Kamera noch eine andere Aufnahmeeinheit, die es ihm erlauben würde, Bezugspunkte und Identifikatoren in der Umgebung ausfindig zu machen; damit kann das „Spotlight Navigation“ anders als die bisher in diesem Kapitel vorgestellten Systeme nicht zur Erweiterung der Realität mit zusätzlichen computergenerierten Inhalten eingesetzt werden.

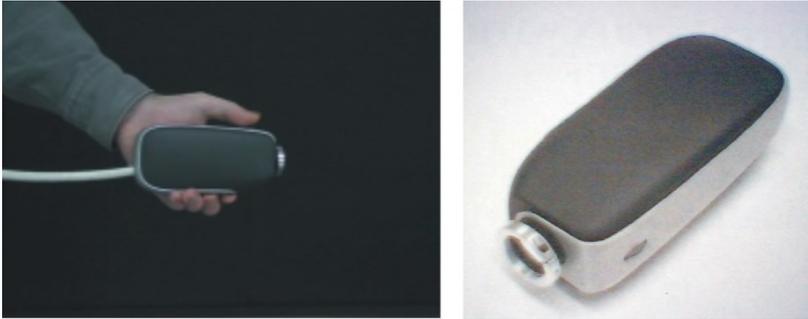


Abbildung 11: Prototyp des "Spotlight Navigation"-Systems [16]

Als tatsächlich neue Technologie ist die für das System entwickelte Benutzerschnittstelle einzustufen: Der schon im Namen vorweg genommenen Metapher der Taschenlampe folgend, bewirkt ein Schwenken des Gerätes, dass andere Bereiche der virtuellen Arbeitsfläche „beleuchtet“ werden; der Benutzer setzt so die angezeigten Bildausschnitte mental zu einem großen Bild zusammen. Die Interaktion mit der Benutzerschnittstelle erfolgt durch die bereits vorgestellte mausartige Handhabung und einem auf dem Projektor angebrachten Zoom-Rad, das das „Hineintauchen“ in Dokumente gestattet: Je weiter herangezoomt wird, desto mehr Details und Interaktionselemente werden dargestellt – und vice versa beim Herauszoomen. Zusätzlich bietet das System noch die Möglichkeit „einzukuppeln“, d.h. durch Schwenken des Gerätes wird nicht der Maus-Zeiger vor virtuell statischem Hintergrund bewegt, sondern die gesamte Oberfläche wird wie ein Fenster in klassischen graphischen Benutzerschnittstellen verschoben [16].

Eine praktische Anwendung des „Spotlight Navigation“ als Alternative zu herkömmlichen PDAs wäre denkbar, ist aber bisher unwahrscheinlich: Die Leuchtkraft des Mikro-Projektors reicht nicht aus, um in normal beleuchteten Räumen vernünftig damit zu arbeiten. Außerdem zwingt die ausschließliche Interaktion über projizierte Bilder den Nutzer dazu, sämtliche angezeigten Daten auch für Dritte einsehbar zu machen. Für die meisten PDA-Nutzer kommt eine Veröffentlichung ihrer Daten allerdings nicht in Frage. Zuletzt ist die Eingabe von Daten durch „Schreiben“ mit dem Gerät auf der Projektionsfläche für den häufigen Gebrauch zu langsam und fehleranfällig [16].

4 Zusammenfassung und Ausblick

Der rasante Fortschritt bei der Leistungsfähigkeit von digitalen Projektoren bei immer kleineren Abmessungen und gleichzeitig sinkenden Preisen der Geräte erschließt den Projektoren völlig neue Anwendungsfelder, die zum Teil noch nicht abzusehen sind. Welches Potenzial in dieser Technologie steckt, kann an der großen Zahl in diesem Bereich aktiver Forschungsteams und fast ebenso vieler Forschungsrichtungen abgelesen werden: Projektoren werden statisch eingesetzt, um konkurrenzlos große Bildschirmflächen zu erzeugen, zur Simulation einer gemeinsamen Zeichnung zweier räumlich getrennter Personen oder als überlegener Ersatz für herkömmliche Monitore; daneben werden fest installierte Einheiten aus Kamera und Projektor zur Erzeugung visueller Umgebungselemente (z.B. um auf Objekte zu zeigen) genutzt oder, um jede beliebige Fläche eines Raumes in einen Bildschirm zu verwandeln. Projektoren sind inzwischen auch klein genug, dass sie auf dem Rücken oder in der Hand mitsamt einem mobilen Rechner transportiert werden und im mobilen Betrieb erweiterte Informationen zu bestimmten Objekten in die Umwelt einblenden können. Neueste Entwicklungen zeigen sogar Möglichkeiten auf, Projektoren als Ergänzung oder vollständigen Ersatz zu PDAs einzusetzen.

Allerdings zeigen die in dieser Arbeit vorgestellten Prototypen, dass bei den meisten Ansätzen noch ein weiter Weg hin zu einem serienreifen Produkt zu überwinden ist. Dies gilt allerdings nicht nur für die projizierten Benutzerschnittstellen, sondern auch für viele verwandte Themengebiete aus dem Bereich neuartiger Benutzerschnittstellen, wie z.B. „augmented reality“ (dt.: „erweiterte Realität“), berührbare Benutzerschnittstellen, Sprachein- und -ausgabe, interaktive Oberflächen, intelligente Kleidung, usw.. Tatsächlich ist die klare Abgrenzung zwischen den Themengebieten bei manchen innovativen Ideen kaum noch möglich und verschwimmt zusehendst; so finden sich Projektoren auch in diversen anderen Artikeln dieses Bandes.

5. Literatur

- [1] A. Majumder, D. Jones, M. McCrory, M. Papka, R. Stevens, Using a camera to capture and correct spatial photometric variation in multi-projector displays. In *IEEE International Workshop on Projector Camera Systems 2003*
- [2] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: Geometrically aware and self-configuring projectors. In *ACM SIGGRAPH 2003 Conference Proceedings*. ACM Press, 2003.
- [3] N. Takao, S. Baker, J. Shi, Steady-state feedback analysis of Tele-Graffiti. In *Proceedings of the IEEE International Workshop on Projector-Camera Systems*, Oktober 2003
- [4] M. Ashdown, P. Robinson, Experiences implementing and using personal projected displays, In *Proc. ICCV Workshop on Projector-Camera Systems (PROCAMS)*, Oktober 2003.
- [5] A. Butz, M. Schneider, M. Spassova. SearchLight – a lightweight search function for pervasive environments. In *Proceedings of Pervasive 2004*.
- [6] S. Borkowski, O. Riff, J. Crowley. Projecting Rectified Images in an augmented environment, In *Proc. ICCV Workshop on Projector-Camera Systems (PROCAMS)*, Oktober 2003.

- [7] Claudio Pinhanez, The Everywhere Displays Projector: A device to create ubiquitous graphical interfaces. *Proc. of Ubiquitous Computing 2001 (UbiComp'01)*, Atlanta, Georgia, September 2001
- [8] C. Pinhanez, F. Kjeldsen, A. Levas, G. Pingali, J. Hartman, M. Podlaseck, V. Kwatra, P. Chou. Transforming surfaces into touch-screens. IBM Research Report RC22273 (W0112-016), December 4, 2001
- [9] C. Pinhanez, R. Kjeldsen, A. Levas, G. Pingali, M. Podlaseck, and N. Sukaviriya. Applications of steerable projector-camera systems. In *Proceedings of the IEEE International Workshop on Projector-Camera Systems at ICCV 2003*, Nice Acropolis, Nizza, Frankreich, 12. Oktober 2003. IEEE Computer Society Press.
- [10] J. Summet, M. Flagg, J. Rehg, G. Corso, G. Abowd. Increasing the usability of virtual rear projection displays. In *Proc. ICCV Workshop on Projector-Camera Systems (PROCAMS)*, Oktober 2003
- [11] Kenji Oka, Yoichi Sato, Hideki Koike. Real-Time Fingertip Tracking and Gesture Recognition. In *IEEE Computer Graphics and Applications 22(6)*: 64-71, 2002
- [12] A. Licsár, T. Szirányi. Dynamic training of hand gesture recognition system. In *ICPR'04*, Cambridge, UK, IEEE & IAPR, Vol.4, 971-974, August 2004
- [13] S. Nayar, H. Peri, M. Grossberg, P. Belhumeur. A projection system with radiometric compensation for screen imperfections. In *Proc. ICCV Workshop on Projector-Camera Systems (PROCAMS)*, Oktober 2003
- [14] T. Karitsuka, K. Sato. A wearable mixed reality with on-board projector. In *ISMAR 2003*
- [15] R. Raskar, P. Beardsley, J. van Baar, Y Wang, P Dietz, J. Lee, D. Leigh, T. Willwacher. RFIG Lamps: Interacting with a self-describing world via photosensing wireless tags and projectors. In *ACM Transactions on Graphics (TOG)*, Vol. 23, Issue 3, pp. 406-415, August 2004
- [16] S. Rapp, G. Michelitsch, M. Osen, J. Williams, M. Barbisch, R. Bohan, Z. Valsen, M. Emele. Spotlight Navigation: Interaction with a handheld projection device. In *Proceedings of Pervasive*, 2004
- [17] Julius Bahr. Interaktive Oberflächen. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München, Feb. 2004. <http://www.hcilab.org/events/ui-update1.0/>
- [18] Marc Breisinger. 3D-Scanning und Rekonstruktion als Benutzerschnittstelle. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München, Feb. 2004. <http://www.hcilab.org/events/ui-update1.0/>
- [19] C. Pinhanez, F. Kjeldsen, A. Levas, G. Pingali, M. Podlaseck, P. Chou. Ubiquitous interactive graphics., In IBM Research Report RC22495 (W0205-143), 22.05.2002

Umgebungsmedien und Stille Technologie

Florian Block

LFE Medieninformatik

Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
block@informatik.uni-muenchen.de

Zusammenfassung Es wird eine Einführung in den Themenbereich der Umgebungsmedien und Stiller Technologie gegeben und Beispielanwendungen vorgestellt. Danach werden Methodiken zum Entwurf und zur Evaluierung solcher Systeme präsentiert. Ein Abschnitt zeigt die Rolle der Datenabstraktion bei Umgebungsmedien. Abschließend werden die zusammengefassten Forschungsergebnisse gegenübergestellt und die Möglichkeiten und ein Ausblick gegeben.

1 Einleitung

Heutzutage existiert eine Fülle verschiedener Interfaces für jegliche Arten von Geräten. Die meisten von ihnen werden entworfen, um dem Benutzer einen möglichst einfachen und zugleich effizienten Zugang zu dem jeweiligen Gerät oder Medium zu verschaffen und dessen optimale Benutzbarkeit zu gewährleisten. Beispiele hierfür sind konkrete und zweckgebundene Schnittstellen an Hausgeräten oder abstrakte Schnittstellenkonzepte am heimischen Personal Computer, die die Benutzung von Betriebssystemen vereinfachen sollen. Auf der Benutzerseite setzt diese Klasse von Benutzerschnittstellen Aktivität voraus. Der Benutzer muss sie bei einem gewissen Grad an Bewusstsein aktiv bedienen. Dabei kann man die Aktionen, die man über ein solche Schnittstelle durchführt in zwei Kategorien einteilen. Zum einen Aktionen, die einen Befehl des Nutzers darstellen, wobei die Reaktion der Schnittstelle auf diese Aktion nur sekundär ist, wie z.B. das Trennen der Internetverbindung, bei der das Interface nur meldet, ob dies erfolgreich passiert ist oder nicht. Auf der anderen Seite ist beim Abrufen von Informationen (z.B. Abrufen von Aktienkursen) die Aktion an sich eher von sekundärer Bedeutung, wohingegen die Reaktion der Schnittstelle entscheidender ist. Dem Nutzer werden die hierbei resultierenden Daten und Informationen über einen gewissen Zeitraum präsentiert.

An letzteren Typ von Interaktion setzt das Prinzip der Umgebungsmedien und der Stillen Technologie an. Anders als bei den bereits konkreten und zweckgebundenen Schnittstellen geht es hierbei darum, dem Benutzer eine Reihe von Informationen über einen längeren Zeitraum subtil, d.h. ohne die Aufmerksamkeit des Benutzers, zugänglich zu machen. Es geht nicht um die exakte Übermittlung harter Fakten, sondern um eine geeignete Abstraktion von Daten, die dem Benutzer ein ungefähres Gefühl über die Charakteristik der zu Grunde liegenden Informationen geben soll. In

[2] verdeutlichen Craig Wisneski et al. das Prinzip anhand der Natur, dass die verschiedensten Sinneseindrücke, wie Gerüche, Lichtverhältnisse, Wind, Luftfeuchtigkeit usw. einen subtilen Status über einen gewissen Zeitverlauf suggerieren. Die Eindrücke helfen uns, Zusammenhänge zu verstehen. Auch im urbanen Bereich lassen sich dafür Beispiele finden. Offene Türen und Lichter in einem Büro informieren uns unterbewusst über die Aktivitäten anderer Leute.

Demnach nutzen ambiente Anzeigen ein breiteres Spektrum an Interaktivität, indem Sie das gesamte physikalische Umfeld mit einbeziehen. In [3] stellen Ishii et al. fest, dass Menschen hoch entwickelte Fähigkeiten haben, mehrere Datenströme gleichzeitig zu verarbeiten. Während eine Informationsquelle die Hauptaufmerksamkeit benötigt, werden andere Quellen gleichzeitig im Hintergrund berücksichtigt. Z. B. können wir durch Anhaltspunkte wie Licht, Temperatur, Geräusche und Luftzüge eine Vorstellung über das Wetter gewinnen. Außerdem können wir die Aktivitäten von Kollegen oder Passanten durch Geräusche in der Umgebung wahrnehmen.

2 Beispiele für Umgebungsmedien

Im Nachfolgenden werden verschiedene Beispiele aus dem Forschungsfeld der Umgebungsmedien und stiller Technologie dargestellt.

2.1 Ambient Room

In [1,3] wird das Konzept AmbientROOM vorgestellt (s. Abb. 1). Die Vision hierbei ist, dass architektonische Umgebungen, in denen wir leben und arbeiten, zukünftig durch eine Vielzahl von Umgebungsmedien erweitert werden und somit eine neue Form von Mensch-Maschine Interaktion erreicht werden kann.

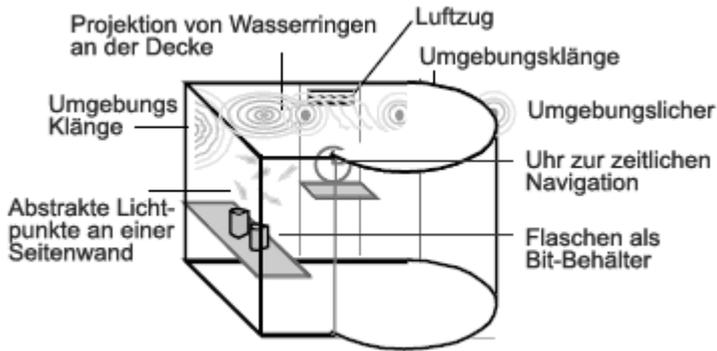


Abb. 1. Aus [3]: Übersicht des AmbientROOM Konzeptes. © 1998 ACM.

In dem Raum befinden sich mehrere Umgebungsmedien wie abstrakte Lichtpunkte an einer der Wände, Umgebungs Klänge oder einer Wasserlampe, die Wasserbrechungen an die Decke wirft (s. Abb. 2). Zudem existieren ein geregelter Luftzug, mehreren Lichtquellen und zwei Steuerelementen in Form einer Uhr und mehrere Flaschen mit Verschlüssen. Diese ermöglichen die eine Interaktion mit den Umgebungsmedien.

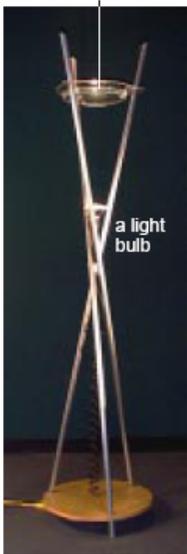


Abb. 2. Waterlamp.
© 1998 ACM [3].

Die Wasserbrechungen werden mit der Aktivität einer vertrauten Person oder Arbeitskollegen verknüpft und geben dem Nutzer somit eine Idee von der aktuellen Tätigkeit dieser Person. Die abstrakten Lichtpunkte werden anhand menschlicher Bewegungen in einem bestimmten Bereich gesteuert. An eine Wand projiziert liefern sie so Aufschluss über den Grad an Beschäftigung im Zielbereich. Eine natürliche Geräuschkulisse, die mit einigen Lichtern im Raum gekoppelt ist, gibt Regen und Vogelgezwitscher in variierender Lautstärke und Modulation aus. Hiermit können Werte wie z.B. die Anzahl ungelesener E-Mail-Nachrichten subtil aber dennoch hörbar suggeriert werden. Kontrolle über das gesamte System bieten als greifbare Objekte die Flaschen und die Uhr. Durch das Öffnen und Schließen verschiedener Flaschen werden soz. Informationen in den Raum freigelassen, also konkrete Datenquellen mit den Umgebungsanzeigen assoziiert. Zudem kann der Benutzer durch das Verstellen der Uhrzeiger den entsprechenden Zeitpunkt herstellen um z.B. nach einer Abwesenheit in Schnelldurchlauf verpasste Eindrücke aufzuholen.

2.2 Waterlamp und Pinweels

Die zwei für sich stehenden Umgebungsanzeigen Waterlamp und Pinweels sind aus dem ambientROOM Projekt entstanden und in [4] beschrieben.

Die Waterlamp (deutsch: Wasserlampe) nutzt einen mit Wasser gefüllten transparenten Schalenbehälter, der von unten durch eine Lichtquelle angestrahlt wird, um Wasserreflexionen an die Decke zu werfen. Die eigentlichen Wellen werden auf der Wasseroberfläche durch drei kleine computergesteuerte Taster generiert, die unabhängig voneinander die Wasseroberfläche berühren können und somit ein charakteristisches Wellenmuster erzeugen. Der gesamte Aufbau ist in Abbildung 2 zu sehen. Wichtig ist, dass die drei Taster unabhängig voneinander angesteuert werden können, wodurch drei unabhängige Datenströme in das Bild der Projektion einfließen können. Die so entstehende Umgebungsanzeige vermittelt dem Benutzer dadurch einzelne Charakteristiken und Zusammenspiele der drei Datenströme.

In dem ambientROOM Projekt wurde herausgefunden, dass der ursprünglich angedachte Luftzug als Umgebungsanzeige sowohl psychisch als auch physisch störend wirkt. Um Luftströme als Metapher dennoch in die Umgebung zu integrieren, wurden Pinweels (deutsch: Windräder, s. Abb. 3) eingesetzt.

Anstatt eines realen Luftstroms werden diese Windräder durch einen computergesteuerten Motor angetrieben. Somit können die einzelnen Räder durch einen entsprechenden „Bitstrom“ angetrieben werden.

Die Anwendung dieser Umgebungsanzeigen ist in verschiedensten Feldern möglich. So können z.B. durch die Wasserlampe Aktivitäten dreier Personen dargestellt werden oder das Zusammenspiel mehrerer astronomischer oder meteorologischer Phänomene. Die Windräder wurden im Beispiel dazu verwendet, den Netzwerkverkehr mehrerer Knoten anzuzeigen. Auch im Bereich der interpersonellen Kommunikation ist der Einsatz dieser Anzeigen denkbar. So könnte die Wasserlampe z.B. die Herzschläge von engen Freunden oder Partnern anzeigen, die durch spezielle Handgelenkuren aufgenommen und gesendet werden.

2.3 LumiTouch

Hierbei handelt es sich um eine semi-umgebende Anzeige, die zur persönlichen Kommunikation zwischen zwei Personen dient. Semi-umgebend meint hierbei, dass im Verhältnis zu den bereits vorgestellten Beispielen mehr Aufmerksamkeit zur Nutzung dieser Schnittstelle aufgebracht werden muss bzw. die Interaktion während der Nutzung von passiv zu aktiv wechseln kann. [5]

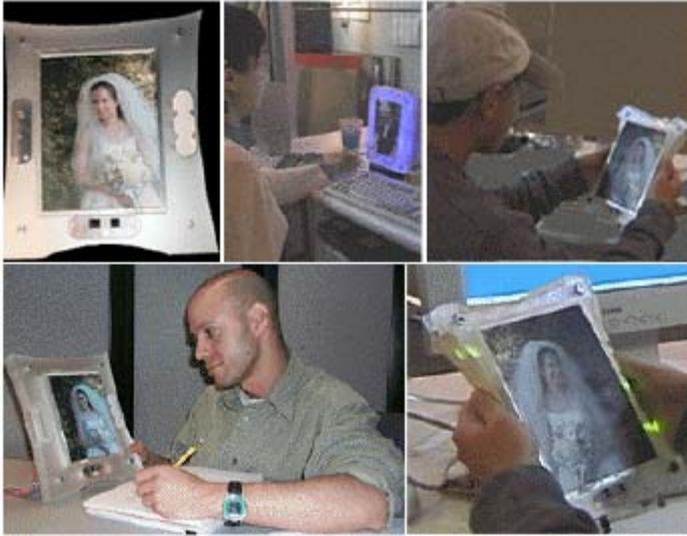


Abb. 3. (Im Uhrzeigersinn, Start links oben) Ein nicht aktiver LumiTouch; ein User, passive Interaktion mit dem LumiTouch; die Rückmeldung des Rahmens auf eine Benutzereingabe im oberen Leuchtbereich; die Anzeige leuchtet, wenn eine Nachricht empfangen wird; ein Benutzer der aktiv mit dem LumiTouch interagiert [5].

Zwei Bilderrahmen mit verschiedenen Sensoren und Anzeigemöglichkeiten (LEDs) interagieren hierbei miteinander. Sitzt ein Benutzer vor dem Rahmen, wird durch ein Glimmen auf der anderen Seite dessen Präsenz angezeigt, bietet so subtile Information über den Aufenthaltsort des anderen und lädt zur aktiveren Interaktion ein. Greift ein Benutzer den Rahmen dann auf, kann er ihn drücken und löst so eine farbliche Veränderung am anderen Ende aus. Dies kann dort auch erwidert werden. Die Dauer und Festigkeit des Drückens beeinflusst die andere Farbgebung (s. Abb. 3). [5]

2.4 HELLO.WALL

HELLO.Wall ist eines der laufenden Beispiele, das im Zuge des von der EU unterstützen Ambient Agoras Projektes entstanden und in [6] beschrieben ist.

Es handelt sich um eine Umgebungsanzeige in der Größe einer Wand, bestehend aus vielen kleineren Lichtzellen (s. Abb. 4). Sie zählt sowohl zu den Umgebungsmedien als auch zu dem Feld der informativen Kunst (s. auch 2.5) und kann auch wie das beschriebene LumiTOUCH zur aktiveren Kommunikation wechseln. Dazu „leiht“ es sich persönliche tragbare Kleincomputer der Nutzer, mit denen es drahtlos kommunizieren kann.



Abb. 4. Interaktion mit der HELLO.WALL durch einen Handheld als „ausgeliehene Anzeige“. [6]

HELLO.WALL unterscheidet bei der Interaktion drei verschiedene Zonen, der Umgebungszone, der Benachrichtigungszone und der Interaktionszone. Je nach Nähe eines Users zu der Wand befindet er sich in einer der drei Zonen. In der Umgebungszone zeigt das Display allgemeine und personenunabhängige Informationen im Sinne der Umgebungsmedien an. In der Benachrichtigungszone kann die Wand spezielle Lichtmuster erzeugen oder auch konkrete Interaktionen über die geliehene Anzeige des tragbaren Computer initiieren. Tritt der Benutzer noch näher an die Wand, befindet er sich in der Interaktionszone und kann über den tragbaren Computer mit einzelnen Lichtzellen oder -bereichen interagieren und somit Informationen abrufen sowie darstellen. Dabei unterstützt HELLO.WALL auch die gleichzeitige aktive Kommunikation mit mehreren Benutzern.

2.5 Informative Art

Als weiterer Repräsentant für informative Kunst und Umgebungsanzeigen soll an dieser Stelle die Arbeit von Lars Erik Holmquist et al. vorgestellt werden [7]. Hier wird Information in künstlerischen Bildern dargestellt. Durch die Auswahl verschiedener Maler-Stile (Piet Mondriaan, Bridget Riley und Andy Warhol) wird konkrete Information in künstlerische Formen gebracht. Die sich ständig aktualisierenden Bilder geben verschiedenste Informationen wieder, wie in Abbildung 5 verdeutlicht wird.

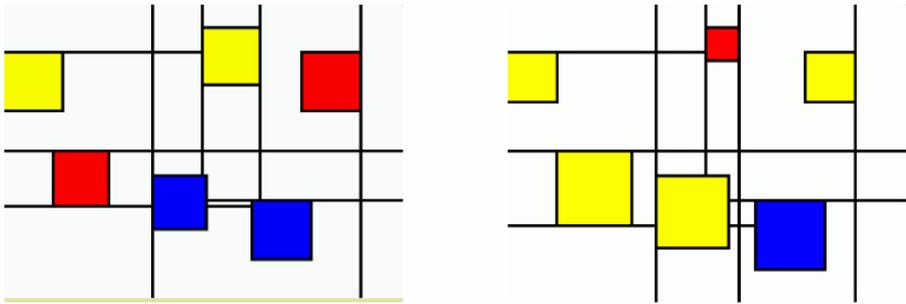


Abb. 5. Das Bild (links und rechts zu zwei verschiedenen Zeitpunkten) im Stil des Künstlers Piet Mondriaan, gibt Auskunft über das Wetter in sechs verschiedenen Großstädten. [7]

2.6 Rückmeldung über Web-Aktivität durch Umgebungsmedien

Die Nutzung von Umgebungsanzeigen im Web-Bereich wird in [8] beschrieben. Hierbei werden Server-Betreiber und –Administratoren über die Art und Intensität der Nutzung ihrer Dienste und Server informiert. Der Einsatz von Umgebungsanzeigen ist hierbei besonders sinnvoll, da er primäre Vordergrundaktivitäten zulässt.

Das fertige System besteht aus mehreren potentiellen Umgebungsanzeigen im Arbeitsbereich, wie z.B. Lampen, einem Tischspringbrunnen und einem Luftbefeuchter, verschiedenen Dienst-Komponenten (in diesem Fall Unterseiten eines Webdienstes) und einer Software, die es erlaubt, die Nutzung verschiedener Komponenten mit beliebigen Anzeigen zu verknüpfen. Hierbei kann auch festgelegt werden, wie die Displays auf die verschiedenen Komponenten reagieren sollen.

Dadurch wird z.B. unter Administratoren oder Mitarbeitern ein permanentes Bewusstsein über den aktuellen Status der Auslastung und Art der Nutzung erreicht, ohne diese von ihren eigentlichen Vordergrundbeschäftigungen abzuhalten.

2.7 Ambient Orb

Die Umgebungskugel [9] ist eine der wenigen kommerziell verfügbaren Umgebungsanzeigen. Für den Heimbereich entworfen, soll er den Bewohnern wichtige Informationen auf subtile Weise übermitteln und gleichzeitig noch als ästhetischer Einrichtungsgegenstand fungieren.



Fig. 6. Ambient Orb. Das leuchtende Ei erzeugt je nach verknüpfter Information verschiedene Farbschattierungen. Copyright by Ambient Devices [9]

Der Besitzer eines Ambient Orbs kann verschiedene Informationen wie Wetterdaten oder Aktienkurse in verschiedene Farbkodierungen umsetzen (s. Abb. 6).

2.8 Geruchs-basierte Ansätze für Umgebungsanzeigen

Die angeführten Beispiele nutzen visuelle und auditive, oder auch taktile Kanäle zur subtilen Informationsübermittlung. Gerade im Bereich der Umgebungsmedien kann man davon ausgehen, dass auch Geruch eine sinnvolle Ergänzung zu den genannten Kanälen darstellen könnte. Im richtigen Leben liefert uns Geruch beiläufig Information über die Qualität von Nahrung, warnt uns vor Bränden oder beeinflusst, wie attraktiv wir eine Person finden. Die Problematik und die Gründe, warum Geruchsschnittstellen bisher so wenig zum Einsatz gekommen ist, wird in [10] erläutert.

Geruch basiert anatomisch auf tausenden verschiedener Geruchsrezeptoren in unserer Nase. Jede dieser einzelnen Rezeptoren spricht auf die charakteristische Schwingungsfrequenz einer chemischen Struktur an. Die Reizung mehrerer dieser Rezeptoren nehmen wir als Gerüche wahr. Verglichen mit dem Sehvermögen, das mit nur roten, grünen und blauen Zapfen und den lichtempfindlichen, jedoch farbenblinden Stäbchen in einem drei-dimensionalen Farbraum zu beschreiben ist, ist der mehrdimensionale Raum der Gerüche äußerst komplex. Dies ist der Grund, warum bisher noch keine systematische und nachvollziehbare Klassifizierung für Gerüche vorhanden ist.

Trotz dieser Problematiken existieren mehrere kommerzielle Systeme zur Computer-gesteuerten Geruchsausgabe und sind von einfachen Desktopvarianten (ca. 50 \$) bis zu groß ausgelegten Lösungen für Auditorien oder Freizeitparks (ca. 25.000 \$) verfügbar. Beim Einsatz dieser Systeme müssen wie bei allen Schnittstellen deren Einschränkungen berücksichtigt werden. Geruchsschnittstellen können aufgrund ihrer Charakteristik nur sich langsam verändernde und lang anhaltende Informationen ausgeben. Zugleich muss man sich auf die Fähigkeiten des Benutzers verlassen, ver-

schiedene Gerüche zu unterscheiden. Untersuchungen haben ergeben, dass man mit wenigen Gerüchen in hoher Qualität und niedriger Intensität die besten Ergebnisse erreicht. Zudem muss man bei der Wahl der Gerüche allergische Reaktionen ausschließen. Bei der Umsetzung der Daten in olfaktorische Signale müssen auch interkulturelle Unterschiede bei der Geruchswahrnehmung beachtet werden. So ist z.B. das amerikanische Rootbeer in Amerika wegen seines eigentümlichen Geschmacks sehr beliebt, in Europa wird es von den Leuten wegen der geruchlichen Ähnlichkeit zu einem in Krankenhäusern verwendeten Desinfektionsmittel abgelehnt.

Zusammenfassend kann gesagt werden, dass Geruchsanzeigen bei dem steigenden Interesse von umgebenden Medien gut geeignet sind und das entwerfen von Prototypen durch die kommerziellen Angebote immer leichter wird.

3 Datenabstraktion und –Wiedergabe bei Umgebungsmedien

Eine gemeinsame Problematik beim Entwurf von Umgebungsmedien ist die Abstraktion der zugrunde liegenden Daten. Dabei treten an zwei verschiedenen Stellen Abstraktion auf. Abstraktion der Datenquelle, d.h., es werden die Signale analysiert und auf gröbere Muster oder Schemata vereinfacht, und Abstraktion bei der Interpretation der daraus entstehenden Informationen.

Diese entstehenden Problematiken werden von dem AROMA-Projekt [12] anhand einer einfachen Demoanwendung verdeutlicht. Durch den Einsatz umgebender Medien soll nahe stehenden Personen auch über große physische Distanz ein gegenseitiges Bewusstsein vermittelt werden. Man soll dabei ein subtiles Gefühl über den Zustand des Anderen bekommen. Dieses erhält man scheinbar mühelos, d.h. ohne Aufwand – allerdings nur, solange nicht etwas passiert, das den Grad unserer Aufmerksamkeit steigern könnte.

Ein Beispielszenario: Zwei Personen, die bisher eng zusammengearbeitet haben, werden räumlich getrennt. Sie entscheiden sich dazu, zu den herkömmlichen Kommunikationsmethoden wie Telefon oder E-Mail einen zusätzlichen Medienplatz auf Ihren Computern einzurichten. Dabei stellt der Medienplatz in Form eines Fensters Aktivitäten des anderen Arbeitspartners dar. Die Darstellung könnte hierbei in Form eines künstlerischen Gemäldes (ähnlich wie in 2.5) oder einer dezenten Klangkulisse realisiert werden, in der das Ändern von visuellen und auditiven Zuständen auf der Gegenseite dargestellt wird. Indem man z.B. die Signale der Mikrofone misst, könnte somit ein Gesamteindruck über die momentanen Aktivitäten gewonnen und beispielsweise in Form von leisem Vogelgezwitscher dargestellt werden.

Hierbei werden die oben genannten Abstraktionsphasen erneut deutlich: Die Signale des Mikrofons werden auf ihre Lautstärke reduziert, d.h. die eigentlich aufgenommenen Daten werden verworfen und fließen nicht in die abstrahierten Daten ein. Danach werden die Lautstärken entsprechend interpretiert und auf der Gegenseite in Vogelgezwitscher verschiedener Intensitäten umgesetzt. Der zweite Grad an Abstraktion steckt hierbei in der Menge an Interpretation, die wir benötigen, um die Anzeige zu lesen.

Da die ursprünglichen, abstrahierten Daten nicht mehr an ein bestimmtes Medium gebunden sind, offeriert dies Spielraum, mit medien-übergreifenden Zuordnungen zu

experimentieren. So kann ein Klangsignal in ein Bild umgesetzt werden oder umgekehrt.

Die semantische Qualität der Anzeige kann dabei durch die Erkennung höherer Elemente im erfassten Datenstrom (z.B. ganze, gesprochene Wörter als höhere Elemente eines Klangsignals) verbessert werden. Dies setzt jedoch einen höheren Anteil an Berechnungsaufwand beim Erfassen der Signale voraus und führt mehr zur genaueren Repräsentation der aufgezeichneten Szenerie, wohingegen der abstraktere Weg mehr zur symbolischen Repräsentation beiträgt.

Die nun aufkommende Frage ist, wo die Vorteile abstrakter Repräsentation gegenüber einer möglichst exakten und informationsreichen Darstellung liegen. Im Speziellen, bezogen auf das AROMA-Projekt und analog zum obigen Szenario erhält Abstraktion die Privatsphäre des jeweils Anderen. Es kann nur gesagt werden, dass der Gegenüber etwas macht, nicht aber was er macht. Im Allgemeinen können folgende Thesen formuliert werden:

- (1) Abstraktion kann informationsreicheren Darstellungen vorgezogen werden indem sie ein peripheres Bewusstsein generieren, dass weniger Aufmerksamkeit benötigt.
- (2) Es ist eine Gefälligkeit gegenüber unserer permanenten Knappheit an Bandbreite (Es wird immer mehr Aktivitäten geben, als wir verarbeiten können).
- (3) Es werden medien-übergreifende Zuordnungen möglich, die es jedem Benutzer ermöglichen das Medium zu wählen, dass für ihn und diesen Zweck am effektivsten ist und allgemein am besten seinen Präferenzen entspricht.

Um die Kernpunkte der Abstraktion anzusprechen und den Entwurf und die Analyse solcher Anzeigen zu systematisieren, wird im Rahmen der AROMA-Projektes folgende generische Architektur vorgestellt:

Bestandteile dieser Architektur ist das *Erfassen* von auditiven und visuellen Daten, deren *Abstraktion* in kompaktere Daten, das *Synthetisieren* von auditiven, visuellen und haptischen Signalen anhand solcher Daten und letztendlich das *Anzeigen* am empfangenden Ende (s. Abb. 7).

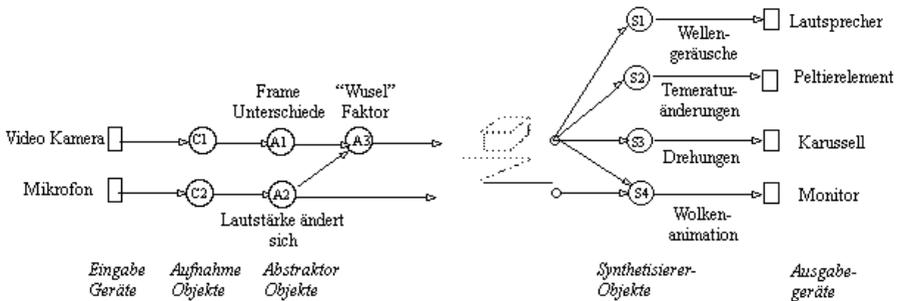


Fig. 7. Architektur des AROMA-Projektes. [12]

Die oben erwähnte, zweite Stufe der Abstraktion wird beim Synthetisieren durchgeführt.

Erfassung

Die Erfassungsseite wird durch das Repertoire an Eingabe- und Aufnahmegeräten sowie Abstraktor-Objekten charakterisiert. Eingabegeräte können Mikrofone, Videokameras oder einzelne Sensoren verschiedener Arten sein. Jedes dieser Eingabegeräte wird durch ein angebundenes Aufnahmegerät angesprochen, das die entsprechenden Signale in einer bestimmten Taktfrequenz abtastet. Zudem haben sie einen Ringpuffer, in dem sie eine Historie der aufgezeichneten Daten abspeichern können. Diese Puffer sind den sog. Abstraktor-Objekten zugänglich. Diese verarbeiten die Signale eines oder mehrerer Aufnahmegeräte durch einfache Vorgänge wie z.B. Akkumulationen oder Verlaufsanalysen und speichern Ihre Ergebnisse wiederum in einen Ringpuffer. Dieser kann unverändert an die Anzeigen geschickt werden oder wiederum als Eingabe für weitere Abstraktor-Objekte dienen.

Kommunikation zwischen Erfassung und Anzeige

Im Allgemeinen ist die Kommunikation zwischen Erfassung und Anzeige nur von geringer Bedeutung. Sie hängt von Faktoren, wie der Distanz von Datenquelle und Anzeige, der Bandbreite der Verbindung, der angestrebten Aktualisierungsfrequenz etc. ab.

Im Bezug auf das AROMA-Projekt werden die gelieferten Daten der Abstraktor-Objekte eingesammelt und über das Netzwerk an die Anzeigen geschickt. Dabei wird bei den Nachrichtentypen zwischen zusammengesetzten Aktivitätsmessungen, visuelle Aktivitätsmessung, Lautsprecher-ID, räumliche Daten und höheren Ereignissen unterschieden. Die Frequenz, in der die Nachrichten ausgeliefert werden, wird so gewählt, dass sie die Abtastzeit sowie die verfügbare Bandbreite der Verbindung berücksichtigt. Eine Ausnahme stellen dabei die höheren Ereignisse dar, die gesendet werden, sobald Sie auf der Abstraktorseite ausgelöst werden.

Anzeige

Eine Anzeige ist durch ihr Repertoire von Ausgabegeräten charakterisiert, die von einer Serie von Synthesizer-Objekten gespeist werden. Mögliche Ausgabegeräte sind Lautsprecher, Bildschirme oder Projektoren und eine breite Auswahl an Wandlern, die haptische oder kinetische Antworten erzeugen.

Die ankommenden Nachrichten werden abhängig vom Nachrichtentyp an einen oder mehrere Synthesizer-Objekte weitergeleitet. Jedes Synthesizer-Objekt ist für die Interpretation einer bestimmten abstrahierten Datenquelle zuständig. Die fertig interpretierte Darstellung gibt es an eines oder mehrere zugeordnete Ausgabegeräte weiter.

Die somit beschriebene Architektur stellt ein Werkzeug zum systematischen Experimentieren und Erforschen von Umgebungsanzeigen und deren zugrunde liegender Abstraktion zur Verfügung.

Die speziellen Ergebnisse des AROMA-Projektes im Rahmen des oben erwähnten Szenarios waren wie folgt:

Das entworfene System machte es in der Tat nach einer kurzen Einlernzeit möglich, ein Bewusstsein entfernter Aktivitäten sogar mit sehr einfachen Anzeigen zu generieren. In der Lernphase betrachteten es die Benutzer als initial schwierig, die Anzeige zu interpretieren und die zugrunde liegenden Informationen zu extrahieren. Ein weiteres Ergebnis war, dass Benutzer selbst Umgebungsanzeigen nicht permanent wahrnehmen. Daraus entsteht die Problematik, dass wichtige Ereignisse unter Umständen verpasst werden. Es wurde daher bei der Interpretation der Daten angedacht, die aktualisierte Standaufnahme der Informationen um eine Historie zu ergänzen (ähnlich der Uhrkontrolle in 2.1). Dies soll dem Benutzer die Möglichkeit geben, verpasste wichtige Ereignisse zu erkennen.

4 Entwurf und Evaluierung von Umgebungsanzeigen

Neben den vielen Beispielen im Bereich der Umgebungsmedien beschäftigen sich nur wenige Arbeiten mit allgemeinen Vorgehensweisen und Methodiken, um Umgebungsanzeigen zu entwerfen und vorhandene Anzeigen zu evaluieren.

4.1 Entwurf von Umgebungsanzeigen

In [11] beschreiben Tobias Skog et al. einige Kriterien, die beim Entwurf von Umgebungsanzeigen beachtet werden sollten. Diese stellen die Ergebnisse aus einer Reihe von Experimenten und Studien dar, die im Rahmen der Evaluierung eigener Umgebungsanzeigen entstanden sind (u.a. auch „Informative Art“, s. 2.5).

Wie bei jeder Informationsanzeige ist der Ausgangspunkt beim Entwurf von Umgebungsanzeigen die Wahl der zu übermittelnden Information. Dabei wird jedoch nicht davon ausgegangen, dass der Benutzer die Informationen aktiv sucht, sondern sie unabhängig davon angeboten bekommt. Es muss damit gerechnet werden, dass die Aufmerksamkeit der Benutzer variiert.

Deswegen ist es in der Entwurfsphase wichtig, einen entsprechenden Geltungsbereich zu bestimmen und die dazu passende Aktualisierungsrate der Anzeige zu wählen. Der Geltungsbereich gibt Aussage darüber, wo und für wen die Anzeige gelten soll. Zweckgebunden soll die Aktualisierungsrate so gewählt werden, dass die Benutzer die Anzeige auf der einen Seite nicht als störend, andererseits aber noch als dynamisch empfinden. Sie hängt zum anderen entscheidend davon ab, wie zeitkritisch die zu Grunde liegenden Informationen sind. Bei sich sehr langsam ändernden Informationen kann der Entwickler dezente Informationen einbinden, die auf die Dynamik der Anzeige schließen lassen.

Bei der Art der Visualisierung gilt es die Balance zwischen Ästhetik und Interpretierbarkeit der Anzeige zu finden. Die Ergebnisse haben gezeigt, dass künstlerische Darstellungen die Eingängigkeit und Erlernbarkeit der Anzeige fördern können. Nutzung von Eselsbrücken (z.B. Landkarten für räumliche Zuordnungen) und Metaphern zeigen ähnliche Effekte.

4.2 Evaluierung von Umgebungsanzeigen

Umgebungsanzeigen sollen dem Benutzer subtil Informationen übermitteln ohne seine Konzentration zu sehr zu stören. Daher ist es schwierig, Effizienz zu definieren und in messbare Ergebnisse zu bringen. Vorhandene Vorgehensweisen wie Langzeitstudien in „belebten Laboren“ sind kosten- und zeitintensiv. Im Gegensatz dazu wird in [12] eine heuristische Evaluation von Umgebungsanzeigen vorgeschlagen. Dabei wird auf vorhandenen heuristischen Evaluierungen traditioneller Anzeigen aufgebaut und die entsprechenden Aspekte umgebender Medien berücksichtigt. Der endgültige Satz von Heuristiken lautet wie folgt (aus [11]):

Passende Wahl von Informationen: Die Anzeige soll ausreichend Informationen bieten. Zu viele Informationen überfüllen die Anzeige, zu wenige lassen den Nutzenansatz außer acht.

Konsistente und intuitive Umsetzung von Informationen: Umgebungsanzeigen sollten nicht überladen sein. Es könnten mehr kognitive Elemente untergebracht werden, wenn der Benutzer sich daran erinnern muss, was die entsprechenden Zustände oder Veränderungen in der Anzeige bedeuten sollen.

System und reale Welt sollen sich entsprechen: Das System sollte die Sprache des Nutzers sprechen und Wörter, Sätze und Konzepte verwenden, die dem Benutzer bekannt sind. Es sollte reellen Konventionen entsprechen und Informationen in einer natürlichen und logischen Reihenfolge präsentieren. Systembasierte Begriffe sollten vermieden werden.

Sichtbarkeit des Zustandes: Es sollte erkennbar sein, in welchen Zustand sich das System befindet. Der Übergang von einem Zustand zu einem anderen sollte klar zu erkennen sein.

Ästhetisches und ansprechendes Design: Das System sollte in dem Umfeld, in dem es eingesetzt wird, ansprechend sein.

Nützliche und relevante Informationen: Die Information sollte in dem Umfeld, in dem sie eingesetzt wird, nützlich und relevant sein.

Benutzerkontrolle und -freiheit: Benutzer können versehentlich Eingaben machen wodurch klar hervorgehobene Funktionen zum rückgängig machen erforderlich werden. Unterstützte rückgängig- und wiederholen-Aktionen.

Einfacher Übergang zu detaillierteren Informationen: Wenn die Anzeige mehrere Level an Informationsdichte anbietet, sollte es die Anzeige dem Nutzer einfach machen die detaillierteren Informationen zu finden.

Nebenläufigkeit der Anzeige: Die Anzeige sollte unaufdringlich sein und in diesem Zustand bleiben bis sie die Aufmerksamkeit des Nutzers braucht. Der Nutzer sollte in der Lage sein, die Anzeige einfach im Auge zu behalten.

Fehlervermeidung: Besser als gute Fehlermeldungen ist ein Design, das Fehler von vornherein ausschließt.

Flexibilität und Effizienz der Nutzung: Beschleuniger – oft von unerfahrenen Benutzern nicht bemerkt – können die Benutzergeschwindigkeit von fortgeschrittenen Benutzern oft erhöhen so dass das System beiden Gruppen eine optimale Bedienung gewährleisten kann.

5 Diskussion

Es existieren viele Beispiele auf dem Feld der Umgebungsmedien. Nur wenige Arbeiten beschäftigten sich mit theoretischen Methoden zum Entwurf, Evaluation und Überprüfung solcher Systeme. Zudem scheint kein tiefenpsychologischer Konsens im Bezug auf die Wirkung und den Einsatz von Umgebungsmedien zu bestehen. Dennoch wird durch einzelne Benutzerstudien das Potential von Umgebungsmedien bestätigt.

Das Spektrum an genutzten Technologien ist sehr breit und bis auf die Geruchs-basierten Anzeigen gut entwickelt. Auffallend ist dabei, dass gerade sehr einfache Technologien effektiv genutzt werden können. Im Hinblick auf die Zukunft wird sich dadurch wohl hauptsächlich die eigene Kreativität limitierend auswirken.

Die Einsatzbereiche von Umgebungsmedien sind hauptsächlich der Arbeitsbereich sowie private Domänen. Auch gesundheitliche Aspekte werden erwähnt.

Mein Vorschlag für einen weiteren, bisher noch kaum abgedeckten Bereich wäre, Umgebungsmedien in Verbindung mit seh- oder geistesbehinderten Menschen einzusetzen. Neben taktilen Informationsträgern könnte man sich zum Beispiel eine auditive und taktile Ergänzung vorstellen, die die Ergebnisse der vorhandenen Forschung im Bereich der Umgebungsmedien sinnvoll in diesem Feld umsetzt. Betrachtet man z.B. das Läuten der Kirchenglocken, war es wohl schon vor hunderten von Jahren einziger zeitlicher Anhaltspunkt für erblindete Menschen. Mit der heutigen hoch entwickelten Technik könnten viel mehr Informationen z.B. per Funk an personifizierte „periphere“ Blinden-Informationssysteme geleitet werden. Da für Sehbehinderte gerade der Gehörsinn zur Orientierung von entscheidender Bedeutung ist, findet das Prinzip der peripheren Informationsvermittlung nicht nur bei taktilen Einflüssen Bedeutung. Auch der primäre auditive Kanal könnte von Umgebungsmedien bereichert werden, ohne störende oder desorientierende Nebeneffekte zu erzeugen. Dasselbe Prinzip könnte auch bei geistesbehinderten Menschen in modifizierter Weise eingesetzt werden.

Literatur

1. Hiroshi Ishii and Brygg Ullmer. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. Proceedings of CHI '97. March 22-27, 1997.
2. Craig Wisneski, Hiroshi Ishii, Andrew Dahley, Matt Gorbet, Scott Brave, Brygg Ullmer, Paul Yarin. Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information. Proceedings of the First International Workshop on Cooperative Buildings (CoBuild '98). February 25-26, 1998.
3. Hiroshi Ishii, Craig Wisneski, Scott Brave, Andrew Dahley, Matt Gorbet, Brygg Ullmer, and Paul Yarin. ambientROOM: Integrating Ambient Media with Architectural Space. Conference Summary of CHI '98. April 18-23, 1998.
4. Andrew Dahley, Craig Wisneski, and Hiroshi Ishii. Water Lamp and Pinwheels: Ambient Projection of Digital Information into Architectural Space. Conference Summary of CHI '98. April 18-23, 1998.
5. Angela Chang, Ben Resner, Brad Koerner, XingChen Wang, Hiroshi Ishii. LumiTouch: An Emotional Communication Device.
6. Thorsten Prante, Richard Stenzel, Carsten Röcker, Norbert Streitz, Carsten Magerkurth. Ambient Agoras – InfoRiver, SIAM, Hello.Wall. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'04), Vienna, Austria. April 24-29, 2004. pp. 763-764.
7. Lars Erik Holmquist and Tobias Skog. Informative Art: Information Visualization in Everyday Environments.
8. Hans-W. Gellersen and Albrecht Schmidt. Look who's visiting: supporting visitor awareness in the web. *Int. J. Human-Computer Studies* (2002) 56.
9. Ambient Orb is a registered trademark of Ambient Devices, see <http://www.ambientdevices.com/cat/contact.html> for details.
10. Joseph Kaye. Culturally Embedded Computing. *Interactions*. january, february 2004.
11. Tobias Skog, Sara Ljungblad and Lars Erik Holmquist. Between Aesthetics and Utility: Designing Ambient Information Visualizations.
12. Elin Rønby Pedersen and Tomas Sokoler. AROMA: Abstract Representation Of Presence Supporting Mutual Awareness. Proceedings of CHI. 1997.

Gefühlsorientierte Benutzerschnittstellen

Dagmar Kern

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`kernd@cip.ifi.lmu.de`

Zusammenfassung In der Mensch-zu-Mensch-Kommunikation spielen Gefühlsäußerungen eine wichtige Rolle. Diese Tatsache soll auch in der Mensch-Maschine-Kommunikation Beachtung finden. Mit Hilfe von gefühlsorientierten Benutzerschnittstellen soll der Computer die Gefühle des Nutzers erkennen und entsprechend darauf reagieren können. Es werden Methoden vorgestellt, die es dem Computer ermöglichen, Rückschlüsse auf die Gefühlslage des Nutzer zu ziehen. Dies geschieht zum einen durch Sensoren, die Körperfunktionen messen oder durch direkte Eingaben, die der Nutzer tätigt. Das Messen der Körperfunktionen alleine reicht nicht aus, es muss eine Zuordnung der Messergebnisse zu den einzelnen Gefühlen geschaffen werden. Ist dieser Vorgang erledigt, können die Erkenntnisse praktisch umgesetzt werden, z.B. in einem Zukunftsauto, welches die Sicherheit des Fahrers erhöht, in einem tragbaren DJ, der die Musik nach dem Gemütszustand auswählt oder in Computerspielen, um diese realistischer zu gestalten.

1 Einleitung

Die Mensch-zu-Mensch-Kommunikation wird häufig durch unsere Gefühle beeinflusst, das liegt zum größten Teil daran, dass unser Gesprächspartner Gefühle erkennen und auch darauf reagieren kann. Warum sollte das eigentlich nicht auch in der Mensch-Maschine-Interaktion möglich sein? Diese Frage führt uns zu den gefühlsorientierten Benutzerschnittstellen.

Wäre es nicht manchmal hilfreich, wenn der Computer oder andere Gegenstände in unserer Umgebung erkennen könnten, wie wir uns fühlen? So könnte unser Zustand verbessert bzw. geändert werden. Wäre es z.B. nicht angenehm, wenn unser Wecker weiß, wie wir uns fühlen und uns dann dementsprechend weckt und nicht wie jeden Morgen mit dem gleichen schrillen Ton? Solch ein Wecker wurde von Wensveen et al [16] entwickelt. Dieser erkennt, mit welchem Gemütszustand eine Person am Abend ins Bett gegangen ist und ermittelt auf dieser Basis die geeignete Weckform für den nächsten Morgen.

Ein anderes denkbare Anwendungsgebiet wäre z.B. der Erhalt von Informationen über die Gefühlslage eines Mitmenschen durch Gegenstände. Einen solchen Gegenstand haben Gaver et al. entworfen, den „Key Table“[3]. Dieser Tisch erkennt, je nachdem, wie jemand nach Betreten der Wohnung seinen Schlüssel, sein Portemon-

naie oder sein Handy darauf ablegt, wie derjenige sich fühlt. Es wurden zwei Mitteilungsmöglichkeiten des Tisches durchdacht: Zum einen, dass je nach dem Gefühlszustand des Eintreffenden ein Cocktail zusammen gemischt wird, mit dessen Hilfe die Gefühle verbessert werden könnten. Die andere Möglichkeit, die als Prototyp umgesetzt wurde, ist ein Bild, das sich je nach Gefühlslage neigt und die Mitbewohner so erkennen können, in welchem Gefühlszustand der Eintretende sich befindet. Eine Skizze des „Key Tables“ ist in Abb. 1 dargestellt.



Abb. 1 „Key Table“: Je nach Intensität des Werfens der Gegenstände auf dem Tisch, neigt sich das Bild, so dass andere Personen ahnen/sehen können, wie der Eintretende sich fühlt [3].

Bisher war es so, dass der Computer die Gefühle des Nutzer vollends ignoriert hat, ähnlich wie ein Gesprächspartner, der in einer Gesprächssituation gar nicht auf die Körpersignale seines Gegenübers eingeht und nur einstudiert antwortet. Diese Reaktion würde die Kommunikation zwischen zwei Menschen vollkommen stören. Beim Computer ist es bisher noch nicht der Fall, weil wir es nicht anders gewohnt sind. Uns ist bewusst, dass der Computer nicht menschlich ist. Dennoch sprechen einige Benutzer häufig mit ihrem Rechner, als ob er eine menschliche Person wäre und bauen eine Art „Beziehung“ zu ihm auf. Wäre es nicht gut, wenn der Computer dann auch dementsprechend antworten würde?

Um eine gefühlsorientierte „Kommunikation“ zur Maschine möglich zu machen, muss zunächst abgrenzt werden, welche Gefühle überhaupt und wie gemessen werden können. Das sind hauptsächlich Gefühle wie Angst, Ärger, Traurigkeit, Freude, Überraschung, Interesse, Desinteresse, Bestürzung, Langeweile und Leid. Indikatoren für diese Gefühle können über Körperfunktionen ermittelt werden, z.B. über den Puls, die Hautleitfähigkeit, die Temperatur, den Blutdruck und andere über Sensoren messbare Funktionen. Wichtig dabei ist, dass eine Zuordnung zwischen den Körperfunktionen und den einzelnen Gefühlen geschaffen werden muss. Wenn die Gefühls-

lage des Benutzers erkannt wird, kann von Seiten des Computers auch entsprechend darauf reagiert und der Nutzer absichtlich beeinflusst werden.

2 Gefühle und Emotionen messen

Idealerweise sollten Computer Gefühle genauso gut wie die Mitmenschen erkennen. Um dies verwirklichen zu können, müssen Möglichkeiten geschaffen werden, Gefühle zu messen. Körperfunktionen, die meist mit den Gefühlen einhergehen, können relativ einfach mit Hilfe von Sensoren überwacht werden. Komplizierter ist die Zuordnung des Gemessenen zu den eigentlichen Gefühlen.

Die Schwierigkeit, Gefühle in Worte zu fassen, hat Jerome Kagan [5] mit der Wettermetapher anschaulich gemacht. Diese besagt:

Der Ausdruck von Emotionen bezogen auf äußere Anreize und der Wechsel, dieser Reize in innerliche Gefühle, ist mit dem Wetter vergleichbar. Das Wetter verändert sich je nach Windgeschwindigkeit, Feuchtigkeit, Temperatur und Luftdruck. Gelegentlich verursacht eine Kombination dieser meteorologischen Phänomene einen Sturm, einen Tornado, einen Blitz oder einen Hurrikan. Das sind Ereignisse, die analog zu den intensiven Gefühlen Freude, Erregung, Empörung und Angst gesehen werden können.

Die Meteorologen haben Wege gefunden, unsere Wetterphänomene recht gut in den Griff zu bekommen, d.h. sie relativ eindeutig der entsprechenden Wetterlage zu zuordnen. Analog kann mit Gefühlen verfahren werden. Die Körperfunktionen (Wetterphänomene) werden mit Sensoren gemessen und müssen den Gefühlen (Wetterlage) zugeordnet werden, um eine Gefühlerkennung möglich zu machen.

Um die Gefühle der Nutzer dem Computer übermitteln zu können, müssen neue Hardware (Sensoren für die Körperfunktionen, Kameras, Stimmerkennung) und neue Software (Algorithmen zur Erkennung des momentanen Zustandes mit Hilfe von gespeicherten Muster) eingesetzt werden.

2.1 Sensoren

Ein wichtiger Bestandteil in gefühlsorientierten Computersystemen sind Sensoren. Sie liefern Informationen über den physischen Zustand des Trägers und erfassen Daten in einer kontinuierlichen Form, ohne den Benutzer in seiner momentanen Tätigkeit zu unterbrechen, d.h. ihn z.B. zu fragen, wie er sich fühlt.

Es werden unterschiedliche Sensoren eingesetzt, zur Messung der Leitfähigkeit der Haut, des Blutdrucks, der Atmung und der Muskelanspannung, die auch häufig kombiniert verwendet werden. Als Beispiel ist in Abb. 2 ein Prototyp der verschiedenen Sensoren vom MIT zu sehen [8].

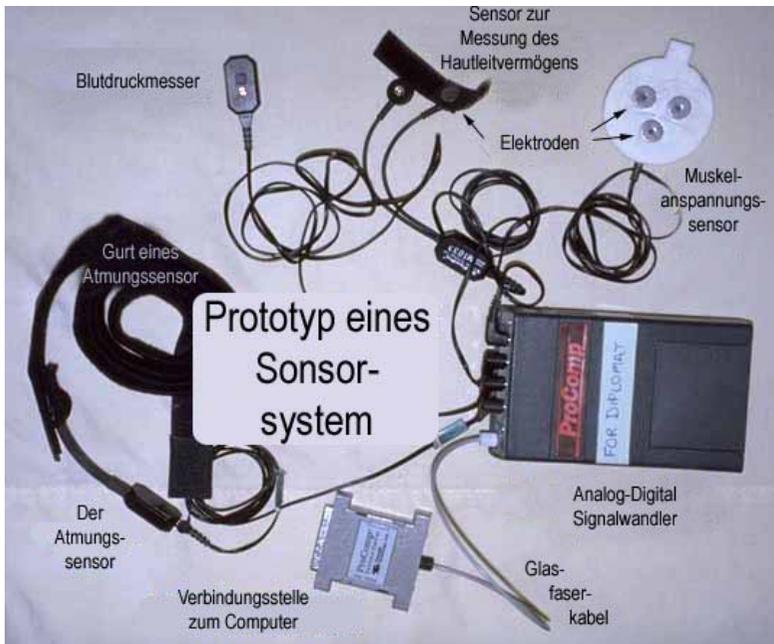


Abb. 2 Prototyp Sensor System vom MIT Media Lab. Das System besteht aus vier Sensoren: Einem Sensor zur Messung des Leitfähigkeit der Haut, einem Blutdrucksensor, einem Sensor zur Messung der Muskelanspannung und einem Atmungssensor [8].

An dieser Stelle soll zur Verdeutlichung wie solche Sensoren funktionieren, beispielhaft das Messen der Leitfähigkeit der Haut beschrieben werden. Bei dieser Methode wird die elektrische Hautleitfähigkeit zwischen zwei Elektroden gemessen. Diese beiden Elektroden sind kleine Metallplatten, die einen sicheren nicht wahrnehmbaren Stromstoß über die Haut legen. Die Elektroden werden typischerweise an den Fingern oder den Zehen angebracht (Manschette), können aber auch an einer anderen Stelle des Körpers getragen werden. Bei der Hautleitfähigkeit werden die Funktion der Schweißdrüsen und die Porengröße betrachtet.

Die Leitfähigkeit ist von einigen Faktoren abhängig: Geschlecht, Ernährung, Hauttyp und, für die gefühlorientierten Schnittstellen wichtig, die momentane Gefühlslage. Die Schweißdrüsen werden durch das sympathische Nervensystem kontrolliert. Das ist der Teil des Nervensystems, der darauf reagiert, ob jemand aufgebracht, z.B. erschrocken oder ängstlich ist. Wenn eine Person erschrickt oder Angst verspürt, erhöht sich die Aktivität der Schweißdrüsen und somit auch die (messbare) Leitfähigkeit der Haut.

Die Sensoren zur Messung der Körperfunktionen müssen leicht zu tragen und robust gegenüber Bewegung sein, um eine Nutzung für den Tragenden attraktiv zu gestalten. Eine Möglichkeit die Sensoren am Nutzer anzubringen ist durch tragbare Computer gegeben, die mit physiologischen Sensoren ausgestattet sind, und so mit Hilfe von gespeicherten Mustern die Gefühlsregungen erkennen und beantworten können. Ein solcher tragbarer Computer ist in Abb. 3 zu sehen. Das BodyMedia SenseWar Arm-

band [2] wurde entwickelt, um eine komfortable nicht störende kontinuierliche Eingabe zu ermöglichen, ohne dass der Nutzer explizit Angaben tätigen muss.



Abb. 3 SenseWar, drahtloses Sensorgerät zur Messung der Körperfunktionen [2].

Als eine andere Möglichkeit der Messung seien die Überlegungen von Wendy Ark et al. in „The Emotion Mouse“ zu erwähnen [1]. Sie verdeutlichen wozu bzw. wie Sensoren im alltäglichen Leben ohne großen Aufwand verwendet werden können. Auf Grund der Tatsache, dass ein Nutzer annähernd 1/3 seiner Computerzeit ein Eingabegerät berührt, sahen sie es als sinnvoll an, eine Maus zu entwickeln, die über Berührung die Körperfunktionen des Nutzers messen kann. Durch die Zuordnung der Körperfunktion zu Gefühlen (siehe Abschnitt 2.4) kann der momentane Zustand des Nutzers ermittelt werden.

2.2 Kameras

Ein weiterer Sensor, der nicht in direktem Kontakt zum Körper stehen muss, ist eine Kamera. Die Kamera filmt während der Computernutzung die Mimik des Nutzers. Dabei werden die Muskelanspannungen des Gesichtes betrachtet, z.B. das Hochziehen der Braue. Aus diesen Informationen können Rückschlüsse auf das momentane Befinden geschlossen werden, in diesem Fall, Erstaunen, Verwunderung. Die Kamera zur Gesichtserkennung kann z.B. in eine Brille (expression glasses) integriert sein, die der Nutzer während der Arbeit am Rechner trägt. Diese Brille macht es möglich Faktoren wie Beleuchtung und Kopfneigen unberücksichtigt zu lassen. Der Nutzer ist in der Lage, sich frei zu bewegen.

Bei der Gefühlsäußerung über die Kamera bleibt es dem Nutzer überlassen, welche Gefühle er äußern möchte und welche nicht, so kann er z.B. falsche Gesichtsausdrücke vorgeben z.B. ein „Pokerface“ aufsetzen, um die wahren Gefühle zu verbergen, genauso wie er es in der Mensch-zu-Mensch-Kommunikation auch machen kann. Möchte er allerdings seine Gefühle äußern, so bietet die Brille einen virtuellen mühelosen Weg dies zu tun. (Details der Brille sind im Paper [14] von Scheier et al nachzulesen).

Für das Erkennen von Gesichtsausdrücken ist eine Vergleichsdatenbank notwendig, in der unterschiedliche mögliche Gesichtsausdrücke gespeichert sind. Die abgefilmte

Mimik wird mit den gesammelten Bildern in der Datenbank verglichen. Bei Untersuchungen von Lisetti wurde beobachtet, dass in einem Gesicht drei Gebiete vorhanden sind, die unabhängig von einander bewegt werden: Der obere Bereich Braue/Stirn, der mittlere Bereich Augen/Lider/Nasenwurzel und der untere Bereich Wangen, Mund, Nase und Kinn. Überraschung ist z.B. meist nur im oberen Teil des Gesichtes zu erkennen durch Falten auf der Stirn, während Freude eher im unteren Bereich zu finden ist, durch Veränderung der Mundwinkel. Lisetti et al haben verschiedene Untersuchungen durchgeführt und festgestellt, dass der Vergleich der einzelnen isolierten Gesichtspartien zu besseren Ergebnissen führt, als ein Vergleich des gesamten Gesichtes [7].

2.3 Gefühle aus der Stimme erkennen

Wer hat seinen Computer noch nicht beschimpft oder wütende Äußerungen gemacht, wenn etwas nicht so funktionierte, wie man es sich vorstellte? Oder wer hat sich noch nicht gefragt, wie sein Telefonpartner sich gerade fühlt?

Das sind unter anderem Gründe, warum das Erkennen der Gefühle aus der Stimme für die gefühlsorientierten Benutzerschnittstellen an Bedeutung gewinnt. Die Vorgehensweise in diesem Bereich ähnelt der Gefühlserkennung mittels Kameras. Die Stimme wird in die unterschiedlichen Bestandteile zerlegt: Tonhöhe, Stimmkraft, spektrale Eigenschaften der Frequenz und zeitliche Eigenschaften (Sprechrates und Pausen). Aus diesen Bestandteilen lassen sich Rückschlüsse auf die Verfassung des Sprechers ziehen. Zum Beispiel ist bei den Gefühlen Ärger, Freude und Angst das sympathische Nervensystem erregt. Das äußert sich dann nicht nur, wie vorher beschrieben, in veränderten Körperfunktionen sondern auch in der Stimme. Sie ist laut, schnell und von starker hoher Frequenzstärke. Wenn das Gegenteil der Fall ist, und eine Person gelangweilt oder traurig ist, wird das parasympathische Nervensystem erregt, was unter anderem zu einer langsamen, niedrigen Tonhöhe mit geringer hoher Frequenzstärke führt. Eine Zuordnung der ermittelten Sprachbestandteile zu Gefühlen wird mit Hilfe von Datenbanken und entsprechenden Algorithmen erreicht.

Untersuchungen haben gezeigt, dass die Genauigkeit des Erkennens der Emotionen aus der Stimme bei folgenden Werten liegt [11]:

Freude	61,4%
Ärger	61,4%
Traurigkeit	68,3%
Angst	49,5%
Normalzustand	66,3%.

Die Erkennung in der Mensch-zu-Mensch-Kommunikation liegt ungefähr im gleichen Bereich, das verdeutlicht, dass auch vom Computer keine 100-prozentige Erkennung erwartet werden muss.

Als typische Einsatzgebiete sind z.B. Call-Center-Anwendungen, Voice Mails oder Telefonunterhaltungen zu nennen.

2.4 Zuordnung von gemessenen Körperfunktionen zu tatsächlichen Gefühlen

Die gemessenen Körperfunktionen sagen in der reinen Form nichts über die tatsächlichen Gefühle aus. Um gefühlsorientierte Schnittstellen entwickeln zu können, muss ein Zusammenhang zwischen den Emotionen und den gemessenen Ergebnissen geschaffen werden.

Eine Möglichkeit die Gefühle den Körperfunktionen zuzuordnen, sind Laboruntersuchungen mit anschließenden Interviews der Versuchspersonen. In Laborsituationen werden Versuchspersonen in bestimmte Gefühlslagen versetzt und die auftretenden Körperfunktionen werden gemessen. Verschiedenen Forschungsgruppen haben mehrere, meist ähnliche Untersuchungen durchgeführt. Hier soll ein Versuch zur Verdeutlichung näher erläutert werden.

In der Untersuchung von Lisetti et al wurden 10 Personen (fünf Frauen, und fünf Männer) einem 35-minütigen Experiment unterzogen. Während des Experiments wurden fünf Gefühle hervorgerufen: Neutral, Angst, Ärger, Traurigkeit und Frustration sowie drei physiologische Signale: Leitfähigkeit der Haut, Herzschlag und Temperatur gemessen. Für das Experiment wurde eine Präsentation entwickelt, diese beginnt mit entspannender Musik und einem entspannendem Naturbild (Ozean). Nach der Entspannungsphase wurde dem Probanden ein Szenario, in Form eines Bildes oder Videoclips präsentiert, welches in der Regel eines der 5 Gefühle hervorruft. Nach dem Hervorrufen des Gefühls folgte wieder eine Ruhephase. Dieser Vorgang wurde so lange wiederholt, bis alle Gefühle einmal erzeugt wurden. Während des Experiments wurden die Versuchspersonen befragt, ob sich das entsprechende Gefühl bei ihnen eingestellt hat. Wenn dies der Fall war, wurden sie nach der Intensität des Gefühls befragt [7].

Zur Zuordnung können verschiedene Algorithmen verwendet werden. Im folgenden soll der „k-nearest Neighbor - Algorithmus“ kurz erläutert werden.

Er berechnet die Distanzvektoren zwischen den Messwerten der Körperfunktionen und den Vergleichsdaten, bei denen die Emotionen bekannt sind. Dann findet er den kleinsten Distanzvektor. Das dazugehörige Gefühl in den Vergleichsdaten stellt die Beziehung zu den gemessenen Daten dar und somit kann der Körperfunktion ein Gefühl zugeordnet werden.

2.5 Probleme beim Messen von Gefühlen

Ähnliche Schwierigkeiten wie bei den Anfängen der Spracherkennung sind auch in diesem Bereich beobachtbar. Erste Untersuchungen bezogen sich nur auf das Erkennen einer Person und waren somit nicht auf andere Nutzer übertragbar. Wie bei der Spracherkennung wird dies auch in diesem Bereich schrittweise verfeinert.

Ein Problem beim Erkennen von Gefühlen über gemessene Körperfunktionen wurde von Rosalinde Picard und ihren Studenten angesprochen. Gefühle äußern sich nicht immer in gleicher Weise, das kann schon bei nur einer einzelnen Person unterschiedlich sein, da sie Tagesform abhängig aber auch durch das Alter und die Kultur be-

dingt sind. Dieses Problem muss bei der Algorithmusfindung, die zur Erkennung der Gefühle eingesetzt wird, berücksichtigt werden [12], [13].

Die Mehrdeutigkeit der mitgeteilten Gefühle stellt ein weiteres Problem dar. Auch in der Mensch-zu-Mensch-Kommunikation versteht nicht jeder Gesprächspartner die Gefühlsäußerungen in gleicher Weise, wie wir es z.B. aus „geschriebenen Gefühlen“ mit Hilfe von Emoticons (Symbole zum Ausdruck von Emotionen durch Tastaturkürzel) in Emails kennen, man kann nie genau sagen, was der Kommunikationspartner nun wirklich mit dem angegebenen „Smile“ meint, man interpretiert seine Wahrnehmung. Die Fähigkeit des Interpretierens ist dem Computer leider nicht gegeben, er benötigt spezielle Algorithmen, die es ihm ermöglichen möglichst zutreffend das derzeitige Gefühl zu erkennen.

2.6 Self-Report

Eine andere Möglichkeit dem Rechner die Gefühle mitzuteilen, ist der „Self-Report“ [13]. Dabei gibt der Nutzer Auskunft über seinen emotionalen Zustand. Dazu könnte z.B. ein Menü oder verschiedene Icons verwendet werden, welche dem Nutzer mögliche Gefühlszustände anbieten, aus denen er sein momentanes Gefühl auswählen und so übermitteln kann. Vorteil dieser Methode ist, dass der Nutzer die volle Kontrolle über die Gefühlsäußerung hat. Er kann entscheiden, was er mitteilen möchte und was nicht.

Nachteil bei dieser Art der Äußerung ist, dass der Nutzer seine Gefühle nicht immer konkret einschätzen kann und die Primärtätigkeit unterbrochen werden muss, um einen Button oder ähnliches zu betätigen. Die Auswahl an angebotenen Gefühlen kann ziemlich groß und unübersichtlich werden, und eventuell trotzdem das gewünschte Gefühl nicht enthalten. Ein großes Problem besteht darin, dass der konzentrierte Nutzer einfach vergisst, den Button zu drücken.

Ein sinnvolles Anwendungsgebiet ist in der Verbesserung von bestehenden Schnittstellen/ Geräten zu finden. Z.B. kann die Nutzbarkeit eines Gerätes durch Abfrage der Gefühle bei der Bedienung vor allem in der Fehlersituation verbessert werden. Der Self-Report ist vor allem dort einsetzbar, wo die Primärtätigkeit sowieso unterbrochen werden muss.

3. Gefühle und Emotionen beantworten

Nachdem der Computer festgestellt hat, wie wir uns fühlen, soll er darauf reagieren, diese Gefühle beeinflussen bzw. neue hervorrufen.

Ein großes Problem dabei ist: Wie soll der Computer überhaupt reagieren, z.B. um Benutzerfrustration zu mindern? Bei der Planung von gefühlsorientierten Schnittstellen muss untersucht werden, wie ein menschlicher Gesprächspartner in einer gegebenen Situation reagiert, es muss versucht werden, diese Art möglichst gut zu imitieren.

Zum Beispiel könnten Aussagen aufgegriffen und vom Computer verwendet werden. Er könnte ein „Gespräch“ zum Stressabbau initiieren „Kein Guter Tag heute, was?“.

Die Gesprächssituation wird realistischer, wenn der Computer in der 1. Person von sich spricht und so eine freundschaftliche Atmosphäre aufbaut.

Er muss allerdings „unterscheiden“, ob ein Nutzer willig ist zu sprechen oder nicht. Im alltäglichen Leben zeigen wir durch aktives Zuhören Empathie und Sympathie, dies muss in der entsprechenden Software ebenfalls umgesetzt werden [13].

Ratschläge stellen eine weitere Möglichkeit der Antwort dar. Dies kann aber auch ins Negative umschlagen, wie es zum Beispiel bei der Büroklammer im Office Paket häufig der Fall ist. Dabei wird aus der Arbeitsweise versucht zuerkennen, was der Nutzer machen möchte, um so hilfreiche Tipps zu geben. Diese Büroklammer strotzt vor Freude und trifft häufig nicht das, was der Nutzer eigentlich will, was bei einigen Anwendern zu einer frustrierten Suche des Ausschaltkopfes führen kann. Dieses Beispiel soll zeigen, dass bei der Erzeugung von Emotionen immer Vorsicht geboten ist, um nicht genau das Gegenteil zu erreichen.

Eine weitere Möglichkeit, Emotionen beim Nutzer hervorzurufen, ist eher implizit. Entsprechend der Gefühlslage führt der Computer Tätigkeiten aus, die für den Nutzer nicht direkt offensichtlich sind, z.B. wird die Musik angepasst oder eine angenehme Atmosphäre durch Dimmen des Lichtes in einem Raum erzeugt.

3 Anwendungsbeispiele

Im folgenden sollen einige praktische Beispiele erläutert werden, die zeigen, dass gefühlsorientierte Benutzerschnittstellen im alltäglichen Leben z.B. beim Autofahren, beim Musikhören oder auch in Computerspielen, ihre Anwendung finden können.

4.1 Das Zukunftsauto

Ein Anwendungsgebiet zum Einsatz von gefühlsorientierten Benutzerschnittstellen ist in der Automobilindustrie zu finden. Die Schnittstellen sollen dazu dienen, das Fahren sicherer zu machen. Aus der körperlichen Verfassung des Fahrers lassen sich Rückschlüsse auf seine Fahrtüchtigkeit ziehen. Ärger, Stress, Müdigkeit, Panik, Frustration und Traurigkeit beeinflussen das Fahrverhalten negativ. Müdigkeit ist einer der gefährlichsten Zustände während des Autofahrens, durch den sogenannten „Sekundenschlaf“ sind schon viele Unfälle verursacht wurden.

Um ein sicherer Fahrer sein zu können, ist es wichtig mehr über die eigenen Gefühle und ihre Handhabung zu wissen. Aus dieser Erkenntnis heraus haben Fatma Nasoz et al. [9] eine Schnittstelle entwickelt, die dem Fahrer hilft, seinen Gefühlszustand besser zu erkennen und damit besser umzugehen. Beim „Multimodal Affective Intelligent Interface“ handelt es sich um eine Schnittstelle, die ihre Informationen mit Hilfe von Sensoren bezieht, die die physiologischen Signale wie die Hautleitfähigkeit, den Herzschlag, und die Temperatur messen, den Gesichtsausdruck beobachten und hö-

ren, wie der Fahrer spricht und somit alle vorher beschriebenen Methoden der Messung verwendet. Die Schnittstelle will auf zwei Arten helfen, das Fahren sicherer zu machen, zum einen werden die Gefühle und der physiologische Zustand erkannt und zum anderen soll der Fahrer direkt gewarnt werden. Dies kann durch eine Meldung erfolgen, wie z.B. „Du bist müde, es ist gefährlich weiterzufahren“. In anderen Situationen z.B. bei Frustration oder Stress kann es hilfreich sein ein Fenster zu öffnen oder beruhigende Musik abzuspielen, dies könnte automatisch angestoßen werden, so dass der Fahrer sich weiter auf den Verkehr konzentrieren kann. Abb. 4 zeigt einen möglichen Aufbau eines Zukunftsautos.

Da diese Art der „Überwachung“ einen großen Eingriff in das Verhalten des Fahrers darstellt und er sich auch massiv dadurch gestört fühlen kann, ist es wichtig zu erwähnen, dass die Messung abschaltbar ist. Dem Fahrer muss es immer freiwillig überlassen sein, ob er sich überwachen lassen möchte oder nicht.

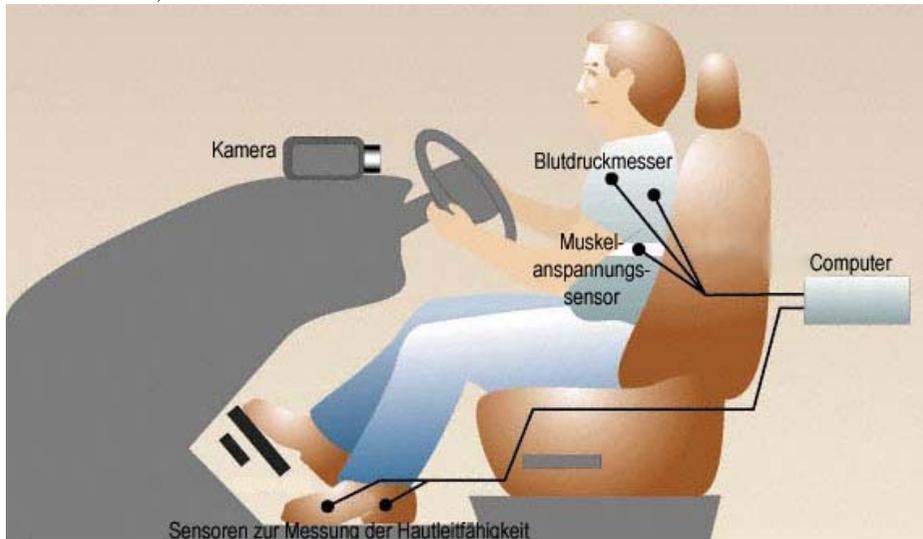


Abb. 4 Versuchsaufbau in einem Zukunftsauto, um mehr über den Stress des Fahrers und andere Gefühlszustände zu lernen, welche die Fahrsicherheit beeinflussen können [13].

Um zu zeigen, dass das Zukunftsauto nicht mehr Zukunft, sondern bereits Gegenwart geworden ist, sei auf den Spiegel-Online Artikel [15] verwiesen. Das dort beschriebene Auto, welches auf dem kommenden Genfer Automobilsalon vorgestellt wird, ändert je nach Gefühlslage des Fahrers das Ambiente des Innenraumes: Ist der Fahrer beispielsweise aggressiv, soll ein sanft gedimmtes Licht helfen, ist er müde ein grelles.

4.2 Steuerung und Erzeugung von Musik

Niemand wird bezweifeln, dass Musik tiefgründige Gefühle hervorrufen kann und starken Einfluss auf menschliche Emotionen hat. Über Messungen von physiologischen Signalen kann die Musik je nach Stimmungslage angepasst werden.

Bei dem Zusammenspiel von gefühlsorientierten Benutzerschnittstellen und Musik muss beachtet werden, dass es zwei Ansätze gibt, zum einen die Steuerung von Musik (vorhandene Datenbank notwendig) und zum andern das Erzeugen von Musik abhängig von der momentanen Verfassung.

Betrachten wir zunächst die Steuerung von Musik. Diese kann zum Beispiel durch den tragbaren DJ [13], von Picard et al entwickelt, realisiert werden. Dieser DJ wählt nicht nur Musik nach der momentanen Hitliste aus, sondern abhängig vom Gefühlszustand des Hörenden. Von früheren Hörgewohnheiten „kennt“ das System die Musik, die der Hörer bevorzugt, wenn er ruhig oder aktiv ist und welche dazu geeignet ist, den Hörer zu beruhigen bzw. aufzumuntern. Für das Abspeichern dieser Hörgewohnheiten sind Trainingseinheiten notwendig, in denen der Computer speichert, wie der Hörer auf das Musikstück reagiert. In späteren Versionen könnte man sich vorstellen, dass dieses Training entfällt und der Nutzer eine im Netz verfügbare Datenbank auf den Computer lädt, die diese Zuordnung bereits gemacht hat (Durchschnittswerte von unterschiedlichen Versuchspersonen). Es besteht die Möglichkeit dem Computer mitzuteilen, welche Stilrichtung man bevorzugt, was er bei der Musikauswahl berücksichtigen soll [4].

Es ist wichtig zu bemerken, dass der Nutzer auch hierbei die volle Kontrolle über das System hat, er kann es nach Belieben an- und ausstellen. Das System wird nicht versuchen einen Hörer aufzumuntern, wenn dieser es nicht direkt oder indirekt befohlen hat. Dem Hörer wird nichts auferlegt, es sei denn, er möchte es so.

Picards Forschungsassistentin Teresa Marrin hat die Ideen des tragbaren DJs im Prototyp Conductor's Jacket [13] umgesetzt. Darunter kann man sich normale Kleidung vorstellen, die den Tragenden nicht in seiner Bewegungsfreiheit einschränkt. Die Jacke vergleicht Muster von Muskelanspannungen und Atmung mit Merkmalen der Musikformen. Sieben Sensoren zur Messung der Muskelanspannung und ein Atmungssensor wurden in die Jacke integriert. In Abb. 5 ist ein schematische Aufbau des Conductor's Jacket dargestellt.

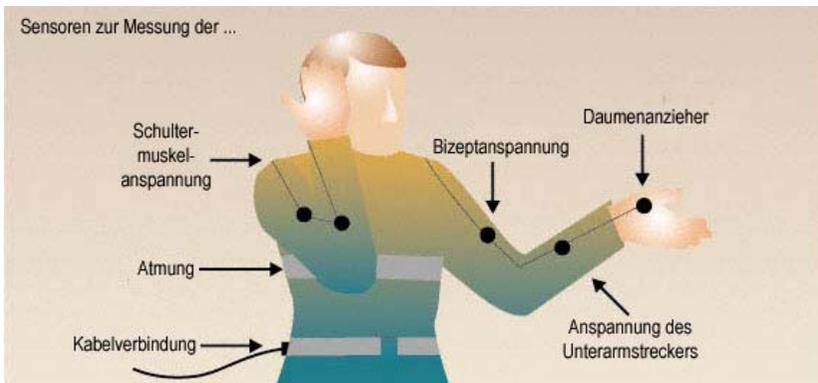


Abb. 5 Schematischer Sensorenaufbau von „Conductor Jacket“ [13].

Einen Schritt weiter geht die Erzeugung von Musik in Abhängigkeit der Gefühlslage. Kim et al. haben festgestellt, dass sich genetische Algorithmen für die Erzeugung von Musik als nützlich erweisen. Die Grundidee von genetischen Algorithmen ist, mit

einem Initialzustand von Lösungskandidaten zu beginnen und darauf aufbauend in zunehmenden Maße bessere Lösungen zu produzieren.

Als ersten Schritt konzentrierten die Forscher sich auf die automatische Generierung von Rhythmen. Diese werden durch vier 16-bit Zeichenketten dargestellt. Um neue Rhythmen zu generieren, entwickelten sie eine Zahl von Veränderungs- und Kreuzoperationen. Zum Beispiel eine Funktion, die zufällig eine Position in zwei Rhythmen auswählt und die Bestandteile (Bits) nach dieser Position tauscht. Diese Funktionen müssen die physiologischen Zustände des Nutzers berücksichtigen, um auf seine Gefühle reagieren zu können. Das macht das Ausbilden des Systems (zur Funktionsgenerierung) durch Versuchspersonen/ Benutzer notwendig. Bei Untersuchungen stellte sich heraus, dass das Erzeugen von Musik mit angenehmen und unangenehmen Rhythmen relativ unkompliziert und schon nach einigen Evolutionsschritten möglich ist [6].

Ein Einsatzgebiet der Musikerzeugung ist z.B. bei Computerspielen zu finden.

4.3 Computerspiele

Computerspiele sind ein wichtiges Gebiet für die gefühlsbezogene Mensch-Maschine-Interaktion. Wenn während eines Spiels die Gefühle des Spielers bekannt sind, können Effekte, wie Erschrecken/ Spannung weiter aufbauen besser erzeugt werden. Zum Beispiel kann wie vorher beschrieben durch Musikgenerierung die Hintergrundmusik dem Gefühl entsprechend verstärkend eingesetzt werden.

Um eine etwas andere Art von gefühlorientierten Schnittstellen kennen zu lernen, soll an dieser Stelle SenseToy von Paiva et al vorgestellt werden [10]. Dabei handelt es sich nicht um die „Eingabe“ der konkreten Gefühle des Nutzers. Stattdessen wird das Verhalten einer Spielfigur durch das Äußern von Emotionen mit Hilfe einer Puppe gesteuert. Bei dieser Puppe handelt es sich um ein mit Sensoren ausgestattetest Eingabegerät, das sechs Gefühlszustände (Freude, Traurigkeit, Ärger, Angst, Überraschung und Schadenfreude wie in Abb. 6 zusehen) übermitteln kann. Je nachdem wie die Puppe bewegt und gehalten wird, werden diese Zustände an das Spiel übermittelt und der Akteur im Spiel reagiert entsprechend. Hierbei handelt es sich um eine greifbare Schnittstelle, mit der Gefühle übermittelt werden können [17,18].

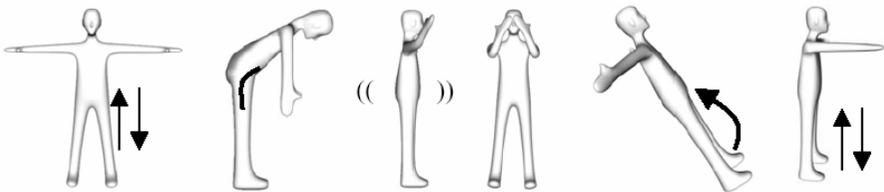


Abb. 6: SenseToy Spieleingabegerät zur Steuerung des Akteurs über Gefühle. Gesten der Gefühlszustände: Freude, Traurigkeit, Ärger, Angst, Überraschung und Schadenfreude [10].

5 Diskussion

Es wurden zwei Methoden zur Übermittlung von Gefühlen des Nutzers an den Computer vorgestellt, zum einen mit Hilfe von Sensoren und zum anderen der Self-Report. Untersuchungen haben gezeigt, dass das Mitteilen von Gefühlen über Sensoren von den Nutzern eher skeptisch angenommen wird, da sie in diesem Fall keine Kontrolle über ihre Eingabe haben, so wie es beim Self-Report möglich ist. Sie haben Angst, zu viel von sich preiszugeben.

Das führt zu dem ethischen Aspekt bei dieser Art der Schnittstelle. Bei den ganzen Betrachtungen darf dieser nicht außer Acht gelassen werden - ob sie für einen selbst sinnvoll erscheinen oder nicht. Emotionen sind absolut persönlich und privat und die meisten Nutzer möchten in der Regel gar nicht, dass ihre Gefühle von einer Maschine erkannt werden. Aus diesem Grund darf bei der Entwicklung solcher Schnittstellen die Rolle des Nutzers nicht unterbewertet werden. Respekt vor den Gefühlen der Personen muss an höchster Stelle stehen. Das führt dazu, dass immer die Möglichkeit gegeben sein muss, die Schnittstelle abzuschalten, wenn der Nutzer sie nicht verwenden möchte. Möchte er allerdings seine Gefühle äußern, ist ihm durch die Schnittstelle eine gute Möglichkeit dazu gegeben.

In dieser Arbeit wurden einige Beispiele zur Kommunikation mit dem Computer beschrieben. Abschließend soll an dieser Stelle noch ein weiteres erwähnt werden, in dem der Computer nur passiv beteiligt ist und somit der ethische Aspekt in den Hintergrund tritt. Das ist das TouchPhone [13]. Bei dieser Entwicklung von Jocelyn Scheirer handelt es sich um ein Telefon, das dem Gegenüber vermittelt, ob der Gesprächspartner den Hörer fest oder locker gedrückt hält, eine Interpretation dieser Information ist dem Kommunikationspartner überlassen, der Computer leitet die Information lediglich weiter. Diese Art der Übertragung kann zur besseren Mensch-zu-Mensch-Kommunikation am Telefon beitragen [19].

Nach den vorherigen Betrachtungen stellt sich dennoch die Frage nach dem Sinn des Ganzen. Die Beispiele haben zwar gezeigt, welche Möglichkeiten vorhanden sind, aber brauchen wir das wirklich? Eine erhöhte Sicherheit im Auto ist gut, aber machen wir uns nicht noch mehr zu Abhängigen der Technikwelt? Will ich nicht lieber meine Musik selber auswählen, als sie mir diktieren zu lassen? Der größte Diskussionsbedarf ist wohl bei den Computerspielen zu sehen: die Spielsituationen werden noch realistischer, die Spielcharaktere noch interessanter, lebendiger. Gleichzeitig besteht aber auch die Gefahr, dass noch mehr Kinder ihre Freizeit mit ihrem Freund Computer verbringen und so die sozialen Beziehungen vernachlässigen.

6 Fazit

Durch gefühlsorientierte Benutzerschnittstellen ist dem Nutzer eine gute Möglichkeit gegeben, seine „Gefühle“ an den Computer zu übermitteln. Mit Hilfe von entsprechenden Algorithmen können die Gefühle von diesem entschlüsselt und entsprechend beantwortet werden.

In dieser Arbeit wurde gezeigt, was bei der Entwicklung von gefühlsorientierten Schnittstellen zu beachten ist. Die Messung der Körperfunktionen, die mit den Gefühlen einhergehen, allein reicht nicht aus, um auf Gefühle reagieren zu können. Wichtig ist die Zuordnung der Körperfunktionen zu den einzelnen Gefühlen. Dies stellt die Hauptschwierigkeit bei der Entwicklung der Schnittstellen dar. Unter anderem müssen Daten von Personen, unterschiedlichen Alters, Geschlecht, Kulturen gesammelt und in Datenbanken zum Vergleich gespeichert werden, es werden Algorithmen benötigt, die die Vergleiche möglichst zutreffend durchführen. Dennoch wird keine 100prozentige Erkennung von Computer erwartet, schon weil dies in der Mensch-zu-Mensch-Kommunikation auch nicht unbedingt gegeben ist.

Es sind einige Beispielszenarios beschrieben worden, das Zukunftsauto, der tragbare DJ und ein mögliches Eingabegerät für Computerspiele. Dies sind Beispiele aus dem täglichen Leben. Interessant sind aber auch die Entwicklungen von Anwendungen für Minderheiten, z.B. für Autisten [13]. Mit einer Jacke, die die Körperfunktionen (Gesten) misst, ähnlich wie beim Conductor's Jacket, wird ihnen die Möglichkeit gegeben, Emotionen mit Hilfe des Computers zu äußern.

Es wurden ausschließlich non-invasive Techniken zur Ermittlung der Gefühle vorgestellt. Es gibt allerdings Entwicklungen, die in Richtung invasive Sensoren gehen, z.B. Implantate, die eingepflanzt werden, um eine ständige Überwachung der „Gefühle“ zu ermöglichen. Die Frage nach der ethischen Problematik wird hierbei allerdings sehr viel bedeutender sein.

Feststeht, dass der Entwicklungs- und Forschungsbedarf in diesem Bereich noch keines Wegs erschöpft ist und wir uns auf weitere Ideen freuen können, die uns unseren Computeralltag „versüßen“ sollen.

Literatur

1. Ark W., Dryer D. C., und Lu D. J. The Emotion Mouse.
In Proceedings of HCI International '99, München, Deutschland, August 1999
2. BodyMedia Inc. www.bodymedia.com (Dezember 2004)
3. Gaver W., Boucher A., Pennington S., Walker B. Weight Workbook
<http://www.equator.ac.uk/PublicationStore/2002-boucher.pdf> (Dezember 2004)
4. Healey J., Picard R., Dabek F. A New Affect-Perceiving Interface and Its Application to Personalized Music Selection.
Workshop on Percetual User Interfaces PUI 98, San Francisco, USA, November 1998
5. Kagan, J. The Nature of the Child. Basic Books, New York, USA, 1984
6. Kim S., André E. Composing Affective Music with a Generate and Sense Approach
In Proceedings of Flairs 2004-Special Track on AI and Music, 2004, AAAI Press
7. Lisetti C. L., Nasoz F. MAUI: Multimodal Affective User Interface.
In Proceedings of the ACM Multimedia International Conference 2002, Juan les Pins, Frankreich, Dezember 2002
8. MIT Media Lab Affective Computing Group, Research on Sensing Human Affect
<http://affect.media.mit.edu/areas.php?id=sensing> (Dezember 2004)
9. Nasoz F., Ozyer O., Lisetti C. L., Finkelstein N. Multimodal Affective Driver Interfaces for Future Cars. In Proceedings of the ACM Multimedia International Conference 2002, Juan les Pins, Frankreich, Dezember 2002).
10. Paiva A., Prada R., Chaves R., Vala M., Bullock A., Andersson G., Höök K. Towards Tangibility in Gameplay: Building a Tangible Affective Interface for a Computer Game
In Proceedings of the 5th international conference on Multimodal interfaces Vancouver, British Columbia, Canada 2003
11. Petrushin V. A. Emotion Recognition in Speech Signal: Experimental Study, Development, and Application.
In Proceedings of the Sixth International Conference on Spoken Language Processing (ICSLP 2000)
12. Picard R.W. Affective Computing: Challenges.
International Journal of Human-Computer Studies Volume 59 , Issue 1-2 (July 2003)
13. Picard R. W. Toward computers that recognize and respond to user emotion.
IBM Systems Journal, Vol. 39, Nos. 3&4, 2000.
14. Schreier J., Fernandez R., and Picard R.W. Expression Glasses: A Wearable Device for Facial Expression Recognition.
CHI '99 Short Papers, Pittsburg, PA (1999)
15. Spiegel Online Lebt das Ding? 20. Dezember 2004, 16:37
<http://www.spiegel.de/auto/aktuell/0,1518,333775,00.html>
16. Wensveen S., Overbeeke K., Djajadiningrat T. Touch me, Hit Me And I Know How You Feel. A Design Approach To Emotionally Rich Interaction.
In Proceeding of ACM Conference on Designing Interactive Systems, Brooklyn, NY, USA, 2000
17. Freund, Stefan. Daten zum Anfassen und Medien zum Berühren. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München, Februar 2004.
<http://www.hcilab.org/events/ui-update1.0/>
18. Ruseva, Radostina. Neue Benutzerschnittstellen für Spiele. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München, Februar. 2004.
<http://www.hcilab.org/events/ui-update1.0/>
19. Schrittenloher, Martin. Neue Formen der entfernten Kommunikation. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München, Februar. 2004.
<http://www.hcilab.org/events/ui-update1.0/>

Benutzerschnittstellen zum Anziehen

Wearability und Fashion

Monika Wnuk

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
wnukm@ifi.lmu.de

Zusammenfassung Durch neue Technologien und Entwicklungsbereiche und die damit zusammenhängende zunehmend drahtlose Übertragung von Informationen, werden Produkte mobiler und tragbarer. Wearable Computing beschäftigt sich mit miniaturisierten, angepassten Benutzerschnittstellen, die in die persönliche Umgebung des Benutzers, also in seine Kleidung oder in einzelne tragbare Elemente eingebettet, also anziehbar sind. Ein anziehbares Objekt, sollte Tragefreundlichkeit - die sog. "Wearability" - aufweisen. Diese Arbeit untersucht und beschreibt die physischen Eigenschaften von anziehbaren, elektronischen Geräten, die nötig sind, um Tragefreundlichkeit zu gewährleisten, um also die bestmögliche Interaktion zwischen dem menschlichen Körper und dem tragbaren Objekt zu ermöglichen. Dabei spielen wohldefinierte Prinzipien, wie die richtige Platzierung oder die gleichmäßige Verteilung des Gewichts von elektronischen Komponenten eine große Rolle. Die Gestalt der anziehbaren Geräte reicht von eingebetteter Elektronik in Kleidungsstücken bis hin zu angepassten Einzelkomponenten, die wie Kleidung genutzt werden, wie z.B. Brillen. Zunehmend wird versucht, diese Äußerungsformen in Hinblick auf gute Wearability zu untersuchen und zu entwickeln, um die Akzeptanz dieser Geräte beim Benutzer und im Markt zu steigern.

1 Einleitung

Der technologische Fortschritt bringt neue Arten der Informationsübertragung mit sich, die es erlauben, mithilfe von drahtloser Kommunikation und immer stärker miniaturisierten, mobileren Endgeräten und alternativen Computerumgebungen unabhängig von einer statischen Desktop-Umgebung auf Informationen zuzugreifen. Neben anderen Entwicklungen, wie beispielsweise den Umgebungsmedien (Ambient Media) [1], steht der persönliche Intimraum des Menschen - der menschliche Körper - als unterstützende Umgebung für Computer im Fokus. Gewöhnliche Laptops oder PDAs jedoch, die eher vom Benutzer gehalten als getragen werden, sind lediglich kleinere Ausgaben von Computern. Man muss seine Tätigkeit unterbrechen, um sie zu bedienen. Durch das bewußte, explizite Benutzen dieser Geräte, können sie nicht durchgehend in Betrieb sein

und observieren, um dem Benutzer situationsbezogen zu helfen. Im Gegensatz dazu bieten anziehbare Benutzerschnittstellen - sog. "Wearable Computer User Interfaces" oder kurz "Wearables" - dem Träger neue, effektivere und mobilere Möglichkeiten des Informationszugriffs.

Ein Wearable ist also ein elektronisches Objekt oder eine Kombination von Eingabe- und Ausgabewerkzeugen, einer Steuerungseinheit (dem eigentlichen Rechner) und einer drahtlosen Kommunikationsvorrichtung, welche in das tägliche Leben, also in die persönliche Umgebung des Benutzers integriert ist und ihm eine operative und interaktive Konsistenz bietet, d.h immer an und erreichbar ist [2]. Dadurch kann ein anziehbares Gerät im Gegensatz zu herkömmlichen Laptops und PDAs die Realität des Benutzers erweitern und, wenn es täglich getragen wird, eine enorme Menge an kontextbezogenen Daten sammeln, um den Benutzer individuell und situationspassend zu informieren. Beispielsweise könnte ein Benutzer auf einer Party einer Frau die Hand schütteln, von der er zwar weiss, dass er sie kennt, aber sich nicht an ihren Namen erinnern kann. Der tragbare Computer tastet den Händedruck ab, startet sein Gesichtserkennungssystem, identifiziert die Dame und zeigt automatisch ihre Kontaktinformation und die letzte eMail an, die der Benutzer von ihr erhalten hat [3]. Wearables haben bereits bedeutend die Arbeit und Leistung von Nutzern beispielsweise in Bereichen der Flugzeugwartung, Routenplanung oder Autoinspektion gefördert.

Da ein anziehbarer Computer an den Menschen gebunden ist, muss seine Gestaltung umso mehr die Dynamik des menschlichen Körpers respektieren. Nachdem nun in diesem Abschnitt in das Thema "Wearable Computing" eingeleitet wurde, kann im nächsten Kapitel in das Konzept und die Prinzipien der Tragefreundlichkeit - der sog. "Wearability" - von tragbaren Produkten tiefer eingestiegen werden. In dem darauffolgenden Kapitel "Äußerungsformen von anziehbaren Computern" werden diese Prinzipien anhand von modischen Beispielen betrachtet und bewertet. Das Unterkapitel "Digitale Accessoires" bietet einen Überblick über die unterschiedlichen elektronischen Komponenten, die in Wearable-Systemen vorkommen. Das Kapitel "Intelligente Kleidung" behandelt dabei Computerumgebungen, die vollständig in Kleidungsstücke integriert sind. Zum Schluss wird in Kapitel "Diskussion und Ausblick" das Konzept der Wearability und die Zukunft von Wearables auf dem Markt kontrovers diskutiert.

2 Tragefreundlichkeit ("Wearability")

Der menschliche Körper ist in Bewegung, seine Form ändert sich permanent und ist verschiedenartig. Ein am Körper getragener Computer, der diese Dynamik respektiert und mit dem Körper in einer aktiven Beziehung steht, entspricht dem Prinzip der Wearability, also Tragefreundlichkeit.

Bisher gibt es im Bereich des Wearable Computing lediglich eine Quelle, die sich mit dem Thema Wearability signifikant und ausführlich befasst - die "Design for Wearability"-Studie der Carnegie Mellon University in Pittsburgh [4]. Die Richtlinien, die im Rahmen der Studie entwickelt wurden, sagen aus, dass tragbare Gegenstände so gestaltet werden sollen, dass sie einer humanistischen

Formsprache entsprechen, d.h. sich an die Form und Bewegung ihres Trägers anpassen, ohne diesen zu behindern. Die Gestaltung von anziehbaren Computern bedarf einer Reihe von Kompromissen zwischen Aspekten der menschlichen Form und Mensch-Maschine-Interaktion und den Zwängen der Technologie und des Nutzzweckes. Im Folgenden werden die wichtigsten Designrichtlinien für die Herstellung von tragbaren Objektformen, vorgestellt.

Platzierung. Die wichtigste Aussage der Studie ist, dass Elemente am Körper ergonomisch und den Träger in seiner Bewegung so wenig wie möglich behindernd platziert werden müssen. Die Kriterien für eine tragefreundliche Platzierung variieren je nach Bedarf an Funktionalität und Erreichbarkeit. Drei Hauptkriterien wurden dabei ermittelt, anhand derer entschieden wurde, welche Stellen am geeignetsten sind. Zum Einen sollen Elemente an Stellen angebracht werden, die in ihrem Umfang bei Erwachsenen in etwa gleich ausfallen, zum Anderen sich relativ unflexibel verhalten, wenn der Körper in Bewegung ist, und die zuletzt eine große Oberfläche besitzen. Diese Kriterien lassen auf die am wenigsten störenden Stellen des menschlichen Körpers schließen, wie in Abbildung 1 zu sehen.

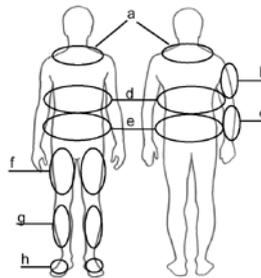


Abbildung 1. Die wesentlichen Stellen, an denen tragbare Objekte angebracht werden sollten: (a) Hals- und Nackengegend (b) Rückseite des Oberarms (c) Unterarm (d) alle Seiten des Brustkorbs (e) Bauch und Hüfte (f) Oberschenkel (g) Schienbein und (h) Fußoberseite (aus [4])

Form. Elemente müssen organisch geformt sein, damit sie der dynamischen Beschaffenheit der menschlichen Figur entsprechen. So muss die Innenseite des Objektes konträr zur Körperoberfläche konkav gestaltet sein, um mit der Körperform direkt abzuschließen. Die Aussenseite sollte hingegen konvex gehalten sein, um störende Ausbeulungen zu vermeiden. Abgeflachte Seiten stabilisieren das Element am Körper und abgerundete Kanten und Ecken machen die Form sicher, weich und tragbar (Abb. 2). Diese Formeigenschaften, die bei keinem anziehbaren Objekt fehlen dürfen, werden in der Studie als humanistische Formsprache beschrieben.

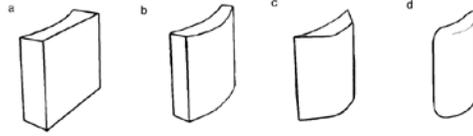


Abbildung 2. Um eine humanistische Formsprache zu erzeugen, muss eine Kombination bestehen aus (a) Konkavität (gegensätzlich zur Körperform), (b) Konvexität, an der Aussenoberfläche des Objekts, (c) Verjüngung, damit das Objekt nicht vom Körper absteht, und (d) Abweichen der Ecken und Kanten (aus [4]).

Bewegungsfreiheit. Die Bewegung des Menschen schränkt die Gestaltung von anziehbaren Objektformen ein, kann aber auch dazu genutzt werden, in Bewegungstests das Profil und die Oberflächenstruktur von Objektformen besser bestimmen zu können. Dies hat das Forschungsteam getan. Zunächst wurden die Elemente analysiert, die eine Körperbewegung ausmachen. Dazu gehören die Mechanik der Gelenke, die Hautdeformation, und die Streckung und Dehnung von Muskeln und Sehnen unter der Haut. Schon durch einfache Bewegungen kann sich die Form des menschlichen Körpers enorm verändern (Abb. 3). Die Freiheit für diese Bewegung kann auf zwei Arten gewährleistet werden. Zum Einen kann man bei der Gestaltung von Objektformen die aktiven Bereiche um die Gelenke herum auslassen. Zum Anderen kann man das Element so gestalten, dass der Körper sich in die Form hinein bewegen kann.



Abbildung 3. Sogar durch kleine Bewegungen verändert sich der Körper erheblich (aus [4]).

Minimaler Körperabstand. Das menschliche Gehirn nimmt eine Art Aura um den Körper herum wahr, welche dazu verwendet werden sollte, um den Abstand des Herausragens von angebrachten Elementen zu bestimmen. Wearables sollten

also laut der "Design for Wearability"-Studie in der intimen Zone des Trägers bleiben, damit sie als Teil des Körpers wahrgenommen werden. Der intime Raum liegt zwischen 0 und 12,5 cm vom Körper entfernt (Abb. 4). Kompromisse sind oft von Nöten, allerdings sollte die Dicke des Elements so stark wie möglich reduziert werden. So werden Komfort und Sicherheit gesteigert, physisch und psychisch.



Abbildung 4. Die Körperaura, die vom Gehirn als zum Körper dazugehörend empfunden wird (aus [4]).

Größenänderung. Körperbau und Zu- und Abnahme von Masse müssen bei der Gestaltung von Objekten berücksichtigt werden. Anziehbare Computer sollten so vielen Körpertypen wie möglich passen. Um die Größenvariabilität zu ermöglichen, kann man sich zum Einen anthropometrische Daten, welche feste Abstände und Längen für verschiedene Körper definieren, anschauen oder Muskel- und Fettzunahme analysieren. Durch die Kombination von starren und flexiblen Teilen können Wearables an diese Variationen angepasst werden. Die flexiblen Formen sollten entweder als Verbindung zwischen den festen Teilen eingesetzt werden, oder die festen Formen in Form von Flügeln erweitern.

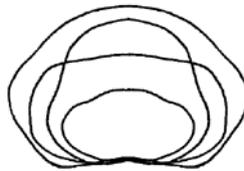


Abbildung 5. Rumpfqerschnitte von verschieden mächtigen Körpern zeigen, wie Körpermaße variieren können (aus [4]).

Anbringung. Elemente können komfortabel angebracht werden, indem sie vielmehr vollständig um den Körper gewickelt werden, anstatt sie mit Gurten oder

Riemen zu befestigen. Zudem ist auch hier wichtig, dass die Anbringung flexibel bezüglich verschiedener Körpermaße ist. Das Design für eine solide, stabile und komfortable Anbringung stützt sich auf Erkenntnisse aus der Bekleidungs- und Outdoor-Industrie. Größenvariable Anbringsysteme können auf zwei Arten realisiert werden. Einerseits durch Einstellbarkeit, z.B. einstellbare Längen von Rucksackriemen, andererseits durch bekannte Größensysteme aus der Bekleidungsindustrie.

Gewicht und Verteilung. Das Gewicht eines tragbaren Computers sollte die Bewegung und das Gleichgewicht des Trägers nicht behindern. Die schwersten Teile sollten dort angebracht werden, wo sich das Hauptgewicht des Körpers befindet - in der Bauch- und Hüftgegend nahe am Körperschwerpunkt - und weg von den Extremitäten. Zudem muss auf eine gleichmäßige Verteilung der Komponenten des Wearables geachtet werden.

Erreichbarkeit. Die Elemente müssen so angebracht sein, dass sie je nach Gebrauchszweck, sei es taktiler, auditiver, oder visueller Natur, am Besten zu erreichen und zu gebrauchen sind. Es macht beispielsweise keinen Sinn, eine visuelle Ausgabestelle unter den Achseln zu platzieren.

Wärmeabstrahlung. Da der Körper atmen muss und empfindlich auf Objekte, die Hitze erzeugen, konzentrieren oder einschließen, reagiert, muss auf eine geeignete Isolierung und Anbringung geachtet werden.

Ästhetik. Ein wichtiger Aspekt von Form und Funktion von anziehbaren Elementen ist ihr Erscheinungsbild. Formen, Farben, Materialien, und Muster werden vom jeweiligen kulturellen und kontextuellen Umfeld bestimmt und diktiert die Gestaltung. Beispielsweise wird ein tragbarer Computer zur Reparatur von Flugzeugen besser vom Träger akzeptiert, wenn er in einem traditionellen Werkzeuggurt aus schweren Leder integriert ist.

Diese Prinzipien basieren auf Erfahrungen des Carnegie Mellon Forschungsteams mit eigens angefertigten, tragbaren Systemen und darauf angewandten Tests, u.a Interviews mit Testbenutzern, sowie auf ausgiebiger Erforschung der Physiologie und Biomechanik des menschlichen Körpers und der Geschichte und Kultur von Bekleidung, Volkstrachten, Schutztextilien und einer Reihe von tragbaren Geräten. Anhand der Forschungsergebnisse und der daraus resultierenden Richtlinien für Wearability wurden dreidimensionale Pads (Polsterkissen) entwickelt (Abb. 6), um weitere umfangreiche Benutzerstudien mit nun in Hinsicht auf Wearability optimierten, tragbaren Objektformen durchzuführen. Die Pads wurden dabei so realitätsnah wie möglich gestaltet, so wurde z.B. darauf geachtet, dass sie dick genug sind, um elektronische Komponenten beherbergen zu können. Die Ergebnisse der Wearability-Tests mit unabhängigen Personen waren größtenteils positiv und besagten, dass die Wahrnehmung von Komfort und Bewegungsfreiheit mit den Objektformen am Körper nahezu nicht beeinträchtigt wurde. Beispielsweise sagte eine Testperson

aus, die Pads seien angenehm und mit wenig körperlicher Einschränkung zu tragen. Somit stellen diese Formen ideale Hüllen für Wearables in Bezug auf dynamische Wearability dar.



Abbildung 6. Dynamische, dreidimensionale Objektformen, die auf Basis der aufgestellten Richtlinien für Wearability von der Forschungsgruppe entwickelt wurden (aus [4]).

3 Äußerungsformen von anziehbaren Computern

Anziehbare Computer können alles sein, angefangen bei kleinen, am Handgelenk getragenen Einzelgeräten bis hin zu komplexen, verteilten tragbaren Systemen, die aus einer Kombination mehrerer Komponenten bestehen. Typischerweise beinhalten diese Systeme eine in einem Gürtel oder Rucksack untergebrachte Rechereinheit, eine am Kopf befestigte Anzeige (ein sogenanntes "Head Mounted Display"), Hardware für drahtlose Kommunikation und eine Eingabeschnittstelle, wie z.B. ein Touchpad. Die Form der Geräte und Komponenten reicht von groß und wuchtig bis hin zu klein und minimal gehalten. Die neue Generation von Wearables gestaltet sich soweit in Accessoires und sogar Textilien integriert, dass sie von normaler Kleidung nicht mehr unterscheidbar ist. Im Folgenden wird die Entwicklung von anziehbaren Benutzerschnittstellen der letzten Jahre anhand einiger Beispielprojekte vorgestellt.

3.1 Digitale Accessoires

Es liegt nahe, elektronische Komponenten in gebräuchliche Accessoires wie Armbänder, Uhren, Brillen, oder Schmuckstücke, beispielsweise Ringe, zu integrieren. Accessoires werden bereits gerne am Körper getragen, sei es zu kulturellen, emotionalen, oder ästhetischen Zwecken. Zudem befinden sie sich von Natur aus

an Stellen des menschlichen Körpers, die gut erreichbar sind, die Bewegung des Menschen minimal behindern und von Körperunterschieden am wenigsten betroffen sind. Accessoires, beispielsweise Uhren, werden ausserdem oft von Natur aus über längere Zeitspannen getragen und eignen sich somit hervorragend für längere, kontinuierliche Informationserfassung bzw. Aktivität.

Digitaler Schmuck In Form von digitalen Schmuckstücken bzw. -sets ist es möglich, Computersysteme physisch in ihre Einzelteile aufzuteilen und so die Komponenten näher an die jeweiligen Ein- und Ausgabe betreffenden Körperteile zu bringen, ohne sich mit langen Kabeln oder anderen Verbindungsstrukturen abfinden zu müssen.

Java Ring. Dallas Semiconductor hat einen elektronischen, Java-basierten Ring (Abb. 7) entwickelt, der ähnlich einer Smart Card automatisch Türen öffnen, den Träger auf Computern einloggen, ihn auf eine sichere Art und Weise identifizieren (Führerschein, Personalausweis), sowie selbstständig Transaktionen im Internet ("One-Button Shopping") durchführen kann [5]. Funktionaler Teil des Rings ist der, als Siegel eingebaute "iButton"-Baustein, auf dem Java-Applets mittels einer virtuellen Java-Maschine ausgeführt werden. Der iButton kann mühelos, mit minimalem Bewegungsaufwand an einen sog. "Blue Dot"-Adapter herangeführt werden, will man beispielsweise eine Sitzung mit einem Computer einleiten. IButtons können in viele verschiedene Dinge, die man nicht zu verlieren gelernt hat, wie Uhren, Brieftaschen, Schlüsselanhänger, etc. integriert werden. In Form des Ringsiegels ist er in eine robuste, praktisch unzerstörbare, auf 10 Jahre Haltbar- und Tragbarkeit getestete Umgebung integriert, ohne schwer oder klobig zu sein. Im Java Ring werden alle denkbaren Identitätsnachweise, die sonst als Vielzahl von Karten oder Buttons in Taschen oder Geldbeuteln verteilt sind, in einer Applikation vereint. Der Träger muss sich nicht auf vorsichtigen Umgang mit dem Gerät konzentrieren, vielmehr verschwindet der iButton ganz im Sinne eines Wearables aus der bewussten Wahrnehmung. Durch seine zudem ästhetische und minimale Ringform ist der Java Ring nicht nur komfortabel und behinderungsfrei zu tragen, sondern dient auch als stylisches Accessoire, auch wenn er für das weibliche Geschlecht etwas zu groß ausfallen mag.

IBM Mobile Phone. Weiterhin hat IBM bereits einen Mobiltelefon-Prototypen entwickelt, der aus verschiedenen digitalen Schmuckstücken besteht (Abb. 8), die miteinander drahtlos, vorraussichtlich über Bluetooth, zusammenarbeiten werden [7]. In Ohringen eingebettete Lautsprecher sollen dabei als Tonausgabe dienen. Sprechen, um sich einerseits mit dem Gesprächspartner zu unterhalten und andererseits das Telefon über Spracheingabe zu steuern, werden die Benutzer in das im Kettenanhänger integrierte Mikrofon. Der Ring ist ausgestattet mit LED-Leuchtdioden, die durch Leuchten einen ankommenden Anruf anzeigen. Dazu gehört optional ein Armband mit einem VGA-Monitor, welcher Name und



Abbildung 7. Java Ring von Dallas Semiconductor (aus [6]).

Telefonnummer von Anrufern anzeigen könnte. Die Schmuckstücke erscheinen vergleichbar mit regulärem Schmuck sehr ästhetisch und "cool", sodass sie durchaus im täglichen Leben tragbar sind, wenn auch nur hauptsächlich von Frauen.



Abbildung 8. Das IBM Mobiltelefon als digitaler Schmuck (aus [7]).

Das Projekt zeigt, wie mobile Geräte und ihre vielfältigen Funktionalitäten vollständig in tägliche Accessoires eingebettet werden können, so dass sie als Teil der normalen Kleidung und nach dem "Design for Wearability" Prinzip des minimalen Körperabstands als zum Körper gehörend wahrgenommen werden. Dadurch kann eine intuitivere und völlig neue Art der Interaktion mit mobilen Geräten geschaffen werden.

Digitale Armbänder Der Arm ist ein beliebter Körperteil für anziehbare Geräte, wie auch in der "Design for Wearability"-Studie festgestellt. Er bietet eine gute taktile Erreichbarkeit, ohne die Bewegung des menschlichen Körpers zu sehr zu behindern und unterscheidet sich in seinem Umfang im Vergleich zu anderen Körperregionen von Mensch zu Mensch relativ wenig. So lassen sich feste Größen für anziehbare Computer entwickeln, die nurmehr durch ein in seiner

Länge variables Halterungsband angepasst werden. Vor allem im Bereich der Wartung und Registrierung von Waren werden am Arm angebrachte Scanner, wie der von Symbol Technologies für UPS entwickelte "Ring Scanner" (Abb. 9 a), verwendet [3]. Die gepolsterten Innenseiten der Arm- und Fingerkomponenten bieten Tragekomfort und verhindern das Wandern der Komponenten um Finger oder Arm. Das leichte und so klein wie möglich gehaltene Utensil wird heute als erleichternde Hilfe in Paketlagern verwendet, eignet sich in seiner Form aber kaum für den privaten oder medizinischen Bereich. Hierzu sind mittlerweile Produkte auf dem Markt, die den Anforderungen für den täglichen Gebrauch entsprechen, wie das entweder am Arm zu tragende, oder zusammengeklappt als reguläres Telefon zu nutzende Mobiltelefon "Tag" von Nec (Abb. 9 b) mit flexibler Tastatur und LED-Anzeige oder das "SenseWear" Armband der Firma BodyMedia (siehe unten), welches Körperwerte messen kann.

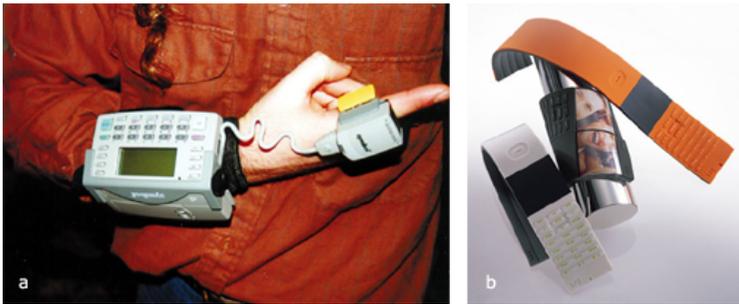


Abbildung 9. (a) Ring Scanner (aus [3]) (b) Flexibles Mobiltelefon "Tag" von Nec (aus [8])

BodyMedia SenseWear Armband. BodyMedia hat ein multifunktionales Sensorsystem für eine erweiterte Körperüberwachung entwickelt, welches, als Armband am hinteren Oberarm getragen, mittels sechs Sensoren permanent den Herzschlag, die Körpertemperatur, die Leitfähigkeit der Haut und die Bewegungen des Patienten misst (Abb. 10). Das System wurde anhand der "Design for Wearability"-Prinzipien und ausgiebiger Usertests so geformt und platziert, dass der Träger es gerne über längere Zeitspannen trägt und dadurch Daten kontinuierlich und genau ausserhalb von Laborumgebungen und Krankenhäusern erfasst werden können.

Die Oberarmgegend eignet sich demnach am besten für das Anbringen des Körperüberwachungssystems [9]. Der Oberarm wird nur selten genutzt für Schmuck oder andere funktionale Accessoires. Zudem ist er geschlechtsneutral, d.h. z.B. im Gegensatz zur Brustgegend (Beispiel Büstenhalter) haben weder Männer noch Frauen Probleme, dort etwas zu tragen. Ein weiterer Punkt ist

der Tragekomfort, den der Oberarm durch seine Weichheit und Gleichförmigkeit ermöglicht und durch seine Struktur und Kraft, könnte sogar ein etwas schwereres Objekt mit wenig Ermüdung getragen werden. Die Oberarmgegend, vor allem der Trizeps, ist eine der am wenigsten hinderlichen Bereiche des Arms und bietet eine gute Möglichkeit, das tragbare, klinische Objekt zu bedecken, um es "privat" zu halten.

Die Form des SenseWear Armbandes ist weich und flexibel gehalten. Symmetrische, flexible Flügel auf jeder Seite des Monitors stabilisieren das Gerät, halten es sicher am Arm, verteilen das Gewicht gleichmäßig, und schaffen einheitlichen Druck auf die Sensoren. Ein elastisches, antiallergisches Band, welches der Befestigung dient, erlaubt dem Bizeps, sich zu dehnen und zusammenzuziehen, ohne zu eng oder zu lose zu wirken. Benutzerstudien mit normalen Personen, Football-Spielern, oder Feuerwehrmännern ergaben eine hohe Akzeptanz des Armbandes und komplette Zuverlässigkeit und Robustheit während verschiedener Alltags-, Trainings- oder Ruhe- und Schlafaktivitäten.



Abbildung 10. "SenseWear" Armband von BodyMedia (aus [9]).

Elektronische Brillen ("Head Mounted Displays") Neben den oben vorgestellten Accessoires spielen "Head Mounted Displays" (HMDs), Computerausgabegeräte, die auf dem Kopf getragen werden und dem Benutzer eine visuelle Ausgabe unmittelbar vor dem Auge bieten, eine große Rolle. Oft werden hierzu Brillen um eine elektronische Vorrichtung erweitert oder bereits als eigenes Gerät entwickelt. Diese Brillen können dann z.B. Orte und Personen wiedererkennen und dem Träger die dazugehörigen Daten anzeigen.

Die ersten HMDs wurden Anfang der achtziger Jahre von Steve Mann am MIT Media Lab entwickelt (Abb. 11 a). Das waren klobige, große Apparate, mit sperrigen 1,5 Zoll-CRTs, welche einen Helm als Halterung erforderten. Lange Antennen für die Funkkommunikation waren daran befestigt. Zusätzlich wurde das HMD kombiniert mit einer 5-Kilo-Ausrüstung, sodass es für den täglichen Gebrauch denkbar ungeeignet war [10]. Gegen Ende der Neunziger Jahre konnten durch den technischen Fortschritt und die Miniaturisierung von

Komponenten Empfang, visuelle sowie auditive Ein- und Ausgabe in minimalen HMDs untergebracht werden, wie Steve Manns "EyeTap" (Abb. 11 d) Projekt [11] oder das Head Mounted Display des Wearable PCs von IBM (siehe unten) zeigt. So werden mittlerweile smarte Anzeig Brillen hergestellt, die entweder kaum unterscheidbar von normalen Brillen oder minimal gehalten als höchst ästhetisches Designobjekt und dadurch durchaus tragbar im täglichen Leben sind.



Abbildung 11. Entwicklung der Head Mounted Displays von Steve Mann angefangen 1980 mit dem (a) Helmdisplay inklusive langer Antennen bis hin zum (d) "EyeTap" Display (aus [10], [12])

IBM HMD. IBM hat einen tragbaren PC mit voll funktionsfähiger "ThinkPad"-Logik entwickelt, der neben einem im Gürtel getragenen PC und einer in der Hand zu haltenden Steuereinheit ein minimales Head Mounted Display (Abb. 12) als Ausgabegerät mit sich bringt [13].



Abbildung 12. Minimales Head Mounted Display des IBM Wearable Thinkpad PCs (aus [14]).

Das Display zeigt ein virtuelles Bild des Computerbildschirms, ohne die Sicht des Trägers zu behindern. Es ist leicht, schmal und ergonomisch geformt, sodass

es auch über längere Zeitspannen getragen werden kann. Zudem gleicht es in seinem Aussehen einem regulären Headset, sodass es auf den ersten Blick nicht auffällt. Ausserdem muss das Auge, wie bei allen HMDs, minimale Strecken zurücklegen, um den gesamten virtuellen Bildschirm zu überblicken. Das Auge ermüdet nicht so schnell, wie bei normalen Monitoren, die bei tragbaren Systemen beispielsweise um den Bauch geschналт oder in Uhrdisplays integriert sind.

3.2 Intelligente Kleidung

Ein weiterer Schritt in der Entwicklung und Gestaltung von anziehbaren Computern ist die direkte Integration von elektronischen Systemen in Textilien. Herkömmliche Vernetzungshardware und starre, am Körper festzuschnellende Komponenten, so klein und portabel sie mittlerweile auch sind, werden ersetzt durch langlebiges, flexibles und sogar waschbares leitfähiges Gewebe und adäquate Ummantelungen von festen Komponenten. Computerbausteine werden in verschiedene Arten von Verschlüssen, wie Reißverschlüsse, Haken, oder Metallknöpfe integriert. Ausgaben werden beispielsweise in Form von "leuchtendem" Garn ermöglicht. Wie bei den elektronischen Accessoires werden Computerumgebungen ermöglicht, die von herkömmlicher Kleidung kaum unterscheidbar sind und sich in höchsten Grade an den menschlichen Körper und seine Bewegung anpassen können. Intelligente Jacken, Unterwäsche, Krawatten oder Schuhe sind da geläufige Äußerungsformen. Forschung und Industrie warten auf mit smarten Hemden, die Herzfrequenz und Atmung messen können (siehe unten), Entertainment-Jacken mit integriertem mp3-Player (siehe unten), Musikjacken mit eingebauter, textiler Tastatur (Abb. 13 a, b), oder Stiefeln, die den Träger automatisch in Chaträumen einloggen (Abb. 13 c).

Computerisierte Kleidung erlaubt es dem Träger, Funktionalitäten moderner Computergeräte im täglichen Leben zu nutzen, während er traditionellen Modetrends entspricht. Der tragbare Computer wiederum kann sich vollständig und ubiquitär in die persönliche Umgebung des Trägers einfügen und sie als Kontext nutzen.

"SmartShirt". Forscher am Georgia Institute of Technology's School of Textile and Fiber Engineering in Atlanta, USA, haben das "SmartShirt" [17] (Abb. 14) entworfen, welches nun auch auf dem freien Markt über die Firma SensaTex erhältlich ist. Das Hemd besteht aus einem leitfähigen Gewebe namens "Silk Organza", einer Mischung aus Seiden- und Kupferfaser. Es kann Herzfrequenz, Körpertemperatur und Atmung messen und den Träger oder Arzt mittels drahtloser Übertragung auf PCs, PDAs, und das Internet oder über Sprachausgabe warnen, wenn ein Problem auftritt. Die Technologie kann eingesetzt werden, um Werte von Astronauten, Leistungssportlern, Feuerwehrleuten, militärischem Personal, chronisch Kranken, oder Kleinkindern zu überwachen.

Im Gegensatz zu herkömmlichen Messungs- und Überwachungssystemen mit

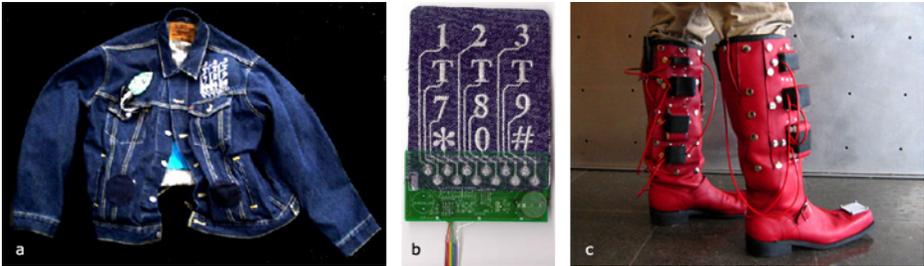


Abbildung 13. (a) Vom MIT Media Lab mithilfe leitfähiger Fasern entwickelte Musikjacke (aus [15]) (b) Textile Tastatur, die in die Musikjacke eingnäht ist (aus [3]). (c) "Seven Mile Boots" (aus [16])

oft klobigen, umständlichen Geräten und Gebundenheit an den Ort der Überwachung ermöglicht die SmartShirt-Technologie, eine für den Träger komplett unsichtbare, komfortable, einfache und behinderungsfreie Überwachung von Körperwerten. Der Träger muss zur drahtlosen Übertragung der gemessenen Werte seinen Standort nicht verändern. Zudem kann die Technologie in jede Art von Material, wie z.B. Baumwolle, Wolle, Lycra, Wolle oder in Stoffgemische eingearbeitet werden ohne das Aussehen, das Tragegefühl, die Flexibilität oder die Beschaffenheit der Textilien, die sie ersetzt zu beeinflussen. Somit bietet das SmartShirt nicht nur hochwertige Funktionalität sondern zugleich ein hohes Maß an Tragefreundlichkeit, wie es herkömmliche Textilien tun.

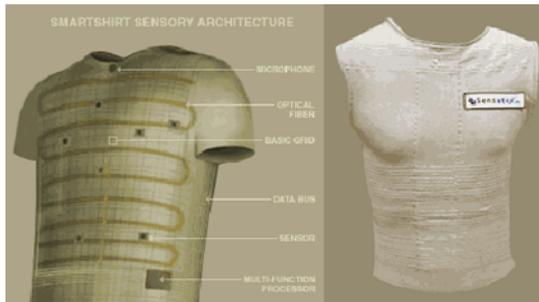


Abbildung 14. "SmartShirt" von SensaTex (aus [18]).

Snowboardjacke "The Hub". Ein weiteres anziehbares System in Form von intelligenter Kleidung ist die von Infineon Technologies in Kooperation mit O'Neill entwickelte Snowboardfunktionsjacke "The Hub" (Abb. 15), mit der man über ein herausnehmbares Chipmodul Musik abspielen und telefonieren kann [19]. Das Modul enthält einen mp3-Player und ein Bluetooth-Modul,

über das ein Mobiltelefon angesteuert werden kann. Über ein flexibles, textiles Bedienelement am linken Ärmel lassen sich mp3-Player und Mobiltelefon bedienen. Zusätzlich wird Sprachsteuerung über das im Kragen eingenahte Mikrofon, über das gesprochen wird, unterstützt. Kopfhörer werden, angeschlossen an das Chipmodul, durch Kabelschlitz und Kabelschleife innerhalb des Jackenrandes zum Kopf geführt. Elektrisch verbunden sind die Einzelteile über flexible Gewebeleitungen.

Die elektronischen Komponenten sind so platziert und flexibel, dass sie den Träger in seiner Bewegung nicht stören und gut erreichbar sind. Das textile Bedienelement ist am linken Ärmel auf Handgelenkhöhe angebracht, sodass es für Rechtshänder gut zu bedienen ist. Es ist stabil genug, um dem Fingerdruck nicht nachzugeben und doch leicht, flexibel und in seiner Form vollständig dem Stoff der Jacke angepasst, sodass es in der bedienungsfreien Zeit als gewöhnlicher Teil des Ärmels wahrgenommen wird. Das Mikrofon in Mundhöhe ist auditiv gut erreichbar. Das Chipmodul hat ein gewisses Gewicht, welches jedoch durch die Platzierung auf Brusthöhe über dem Körperschwerpunkt relativ unmerklich ist. Im Gegensatz zu einem herkömmlichen mp3-Player, der für gewöhnlich in Jackentaschen untergebracht wird, hängt die Jacke nicht zur einen Seite über, sondern behält ihre gleichmäßige Position. Trotz des Gewichts des Chipmoduls ist die Elektronik leicht und behindert nicht den Tragekomfort und die Bewegung. Zwar sind die Kopfhörer nicht physikalisch in die Jacke eingearbeitet, die Kabelführung durch Schlitz und Schleife verhindert aber Störungen durch Kabel und führt das Kabel zum Kopf. Durch die auf Schnee-, Regen-, und Frostfestigkeit getesteten, mechanisch robusten und flexiblen Leitergewebe ist die Jacke waschmaschinentauglich und bügelfest wie herkömmlicher Stoff.

Die Jacke ist ein gutes Beispiel für die Integration eines Unterhaltungssystems in tägliche Kleidung. Weiterhin ist denkbar, dass die Kopfhörer in textiler Form implementiert und die Stromversorgung komplett vom menschlichen Körper übernommen wird.



Abbildung 15. "The Hub" Snowboardjacke von Infineon Technologies in Kooperation mit O'Neill (aus [19]).

4 Diskussion und Ausblick

Bisher waren die meisten anziehbaren Computerprodukte zu teuer für Privatpersonen und wurden vorwiegend für vertikale Nischen hergestellt. So werden sie während der Verrichtung von berufsspezifischen Tätigkeiten getragen, für die ohnehin Ausrüstung benötigt wird, wie z.B. bei der Fluginspektion. Dort haben sich Wearables durch ihre ergonomischen und funktionalen Vorteile gegenüber herkömmlichen, in der Hand zu haltenden Geräten weitgehend durchgesetzt. Der Ringscanner ist hierfür ein gutes Beispiel.

In der Verbreitung von anziehbaren Computern wird nun der Massenmarkt angesteuert. Einige Forscher sehen Wearables sich in jedem Bereich des Lebens etablieren. Andere glauben, dass ihre Zukunft stark von der Akzeptanz in der Gesellschaft abhängt. Dafür ist die Tragefreundlichkeit der tragbaren Produkte ausschlaggebend. So ist noch viel Arbeit im Bereich der Forschung, Entwicklung und des Testens, und das nicht nur in Laborumgebungen, sondern auch im täglichen Gebrauch, nötig.

Beispielsweise arbeiten Designer und Ingenieure bereits an der Reduzierung störender Einflüsse, wie Hitzeausstrahlung und Elektrosmog, indem sie Gewebe verarbeiten, die elektromagnetische Strahlung abschirmen [20]. Zudem wird versucht eine vollständig integrierte Stromerzeugung zu ermöglichen. Infineon Technologies hat hierzu bereits einen Thermogenerator entwickelt, der aus den Temperaturdifferenzen zwischen Körper und Kleidung elektrischen Strom erzeugt [20]. Denkbar wird hier auch sein, Energie über einen Federmechanismus zu beziehen, der 6-8 Watt liefern könnte. Dies ist genug um beispielsweise ein PDA oder Mobiltelefon zu betreiben.

Um Wearables im Massenmarkt zu etablieren, reicht eine gute Wearability der Produkte allein jedoch nicht aus. Ihre Durchsetzung hängt zudem stark von einer guten Marketingstrategie, und einem vernünftigen Absatzort ab. Werden Schmucktelefone nun in Schmuckläden angeboten, oder doch in Mobilfunkläden? Das Scheitern des Schmucktelefons Xelibri von Siemens im Jahr 2003 [21] war beispielsweise darauf zurückzuführen, dass in Modeboutiquen kein geschultes Personal zur Verfügung stand, um Kunden die Technik nahezubringen, und in Mobilfunkläden das Vertrauen in einen guten Absatz des Produkts zu gering war. Zudem reichen Tragekomfort und gutes Zusammenspiel mit dem menschlichen Körper nicht aus, wenn das Gerät keine Verbesserung bzw. Alternative zu herkömmlichen Produkten in ihrer Funktionalität und Zeitersparnis aufweist. So werden Benutzer beispielsweise in der Lage sein müssen, ihre Desktop PCs verlassen und ihre täglichen Aufgaben weiterverfolgen zu können, während sie verbunden bleiben mit EDV und Kommunikationsressourcen [3]. Ist das Aussehen zudem den gesellschaftlichen Vorstellungen und Vorlieben nicht entsprechend, also beispielsweise "uncool", wird das Produkt trotz guter Wearability nicht im Massenmarkt boomen.

Weiterhin wird seit einiger Zeit das Problem mit der Verletzung der Privatsphäre und des Datenschutzrechts von Personen in der unmittelbaren Kommunikationsumgebung des Trägers diskutiert [22]. Wearables könnten ungehindert Personendaten sammeln, wie im Gesichtserkennungsbeispiel aus Kapitel 1 beschrieben.

Auf der anderen Seite werden die Preise mit weiterer Verbreitung von Wearables sinken. Funktionalitäten und Zeitersparnis für den Nutzer werden mit der raschen technologischen Forschung und Entwicklung hinzukommen. So wird das Kosten-Nutzen Verhältnis kontinuierlich verbessert. Zudem werden anziehbare Geräte immer tragfreundlicher und, durch die stetige Steigerung des Mobilitätsbedarfs, attraktiver. Schon jetzt ist zu beobachten, dass Menschen dazu tendieren, mobile Geräte zum Computer zum Anziehen umzufunktionieren. Beispielsweise tragen Leute ihre Handys an Bändern um den Hals. Wird der Weg in die Medien und die Geschäfte durch gute Marketingstrategien geebnet, werden Wearables vielleicht sogar schon in naher Zukunft im Massenmarkt boomen.

Literatur

1. Block, F.: Umgebungsmedien und stille Technologie. Seminar UI-Update 1.0. Ludwig-Maximilians-Universität München. <http://www.hcilab.org/events/ui-update1.0/> (Feb. 2005)
2. Mann, S.: Wearable Computing as means for Personal Empowerment. In: ISWC '98: Proceedings of the 2nd IEEE International Symposium on Wearable Computers, IEEE Computer Society (1998)
3. Billinghamurst, M., Starner, T.: Wearable Devices: New Ways to Manage Information. *Computer* **32** (1999) 57–64 Band 32, Ausgabe 1.
4. Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M., Martin, R.: Design for Wearability. In: ISWC '98: Proceedings of the 2nd IEEE International Symposium on Wearable Computers, IEEE Computer Society (1998) 116
5. Cameron, C., Dayn, B.: Knuckletop Computing: The Java Ring. <http://java.sun.com/features/1998/03/rings.html> (Mär. 1998)
6. Bonsor, K.: How Digital Jewelry Will Work - The Java Ring. <http://electronics.howstuffworks.com/digital-jewelry3.htm> (aufgerufen im Jan. 2005)
7. Bonsor, K.: How Digital Jewelry Will Work - Give Me a Ring. <http://electronics.howstuffworks.com/digital-jewelry1.htm> (aufgerufen im Jan. 2005)
8. Nec Corporation: Tag: Soft-shell mobile phone. <http://www.nec-design.co.jp/showcase/> (2003/2004)
9. Why the Upper Arm? Factors Contributing to the Design of an Accurate and Comfortable, Wearable Body Monitor, erreichbar über www.BodyMedia.de (2002)
10. Mann, S.: Smart Clothing: The Wearable Computer and WearCam. *Personal Technologies* (Mär. 1997) Band 1, Ausgabe 1.
11. EyeTap Personal Imaging Lab: EyeTap: The Eye Itself as Display and Camera. <http://eyetap.org/research/eyetap.html> (2004)
12. Mann, S.: Continuous lifelong capture of personal experience with EyeTap. In: CARPE'04: Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences, New York, NY, USA, ACM Press (2004) 1–21
13. Wilcox, J.: More Details Emerge on IBM's Wearable PC. <http://news.com.com/2100-1040-234689.html?legacy=cnet> (1999)
14. IBM: IBM Research. <http://www-5.ibm.com/se/news/archive/images/computers/wearable-pc/wearable-pc-closseup-guy.jpg> (aufgerufen im Jan. 2005)
15. Bonsor, K.: How Computerized Clothing Will Work. <http://computer.howstuffworks.com/computer-clothing.htm> (aufgerufen im Jan. 2005)

16. Beloff, L., Berger, E., Pichlmair, M.: Seven Mile Boots. <http://randomseed.org/sevenmileboots/> (2003/2004)
17. Bowie, L.: Smart Shirt Moves from Research to Market - Goal is to Ease Healthcare Monitoring. Georgia Institute of Technology, erreichbar über www.gatech.edu (Jun. 2000)
18. SensaTex Inc.: Welcome to SensaTex. <http://www.sensatex.com/index.html> (2002)
19. Schoenrock, R.: O'Neill Unveils Snowboard Jacket with Integrated Infineon Electronics - Entry into Future Market of 'Wearable Electronics'. Infineon News Releases (Jan. 2004)
20. Saße, D.: Europas neue Kleider - Technik wird tragbar. Computer und Technik. <http://www.heise.de/ct/03/04/022/> (Apr. 2003)
21. Longino, C.: Axe Falls on Xelibri. <http://www.thefeature.com/article?articleid=100665&threshold=-1&ref=5270724> (Mai 2004)
22. Wood, C.: Privacy and the Wearable Computer. PC Magazine. http://www.pcmag.com/print_article/0,3048,a%3D20006,00.asp (2001)

Neue Formen der entfernten Kommunikation

Martin Schrittenloher

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
martin@schrittenloher.de

Zusammenfassung Entfernte Kommunikation umfasst heutzutage ein weites Feld von verbalen und nonverbalen Möglichkeiten. Methoden wie Mobiltelefonie, SMS, Email, Chat oder Videokonferenz bieten bereits qualitativ hochwertige Mittel hierfür. Forschungsarbeiten in dieser Richtung haben alle gemeinsam, dass sie versuchen, diese herkömmlichen Wege der Kommunikation zu erweitern. Die auditiven und visuellen Kanäle sollen durch einen für die interpersonelle Kommunikation selbstverständlichen Kanal, namentlich die Berührung oder durch physikalische Einflussnahme entweder erweitert, oder gar vollends ersetzt werden. Aufgrund der Komplexität menschlicher Kommunikation wird die Fülle und Vielfalt der Möglichkeiten zwar betrachtet, der Schwerpunkt dieses Manuskripts liegt jedoch in der taktilen Verständigung. Diese Arbeit setzt sich zum Ziel, anhand einiger Beispielarbeiten herauszustellen auf welchem Stand die Forschung momentan ist und welches theoretische Potential diese Forschungsrichtung für die Zukunft hat. Dafür werden veröffentlichte Artikel vorgestellt und in einem abschließenden Vergleich diskutiert.

1. Einleitung

Möchte man in diesen Tagen mit einer Person kommunizieren, die sich räumlich nicht am gleichen Ort wie man selbst aufhält, so bieten sich eine Vielzahl von Möglichkeiten an, dies umzusetzen. Hier seien mit SMS-Kurznachrichten, Telefonaten über das Mobil- oder Heimtelefon, der wohl häufiger in geschäftlichen Belangen genutzten Videokonferenz, dem Chat oder der Email nur die am weitest verbreiteten genannt. Doch im Vergleich zur zwischenmenschlichen, räumlich nicht getrennten Kommunikation wird deutlich, dass hierbei bislang auf wichtige Aspekte verzichtet werden musste. Die zwischenmenschliche Kommunikation läuft über mehrere Kanäle ab. „Gemeint sind jene Sinnesmodalitäten, mithilfe derer und über die wir unsere Kommunikationspartner wahrnehmen.“ (Pürer, S. 63) Man unterscheidet hier zwischen auditivem, visuellem, taktilem, thermalem, olfaktorischem und gustatorischem Kanal (vgl. Pürer, S. 63).

So sind Berührung und physikalische Einflussnahme im echten Leben ein grundlegender Bestandteil zum Verstehen und Beeinflussen unserer Umwelt.

Die Mobilfunktelefonie konzentriert sich auf den auditiven, die Techniken zur Videokonferenz auf eine Mischung aus ebendiesen und visuellem Kanal. Da auch bei Email-Kontakt oder Chat meist nur der visuelle Kanal angesprochen wird (von weni-

gen Ausnahmen einmal abgesehen) bleiben viele der Kanäle im Bereich der Technik zur entfernten Kommunikation ungenützt.

Durch die schnellen Entwicklungsschritte drahtloser und anderer Kommunikationstechnologien wird es aber zusehends realistischer, Berührungen zu übertragen und Nähe über beliebige Entfernungen hinweg darzustellen.

Zu diesem Thema gibt es viele Arbeiten von Wissenschaftlern, die sich spielerisch, aber stets mit ernsthaftem Gedanken, mögliche Anwendungen ausgedacht, implementiert und getestet haben und teilweise mit erstaunlichen Ergebnissen aufwarten. Um nun eine Diskussionsgrundlage dafür zu schaffen, welche Entwicklungen sinnvoll sind und wo eine Nachfrage absehbar sein wird, ziehe ich als Beispielarbeiten „The Hug -Die Umarmung“, „Tangible Interfaces for Remote Collaboration and Communication - Berührungssensitive Schnittstellen zur entfernten Kommunikation am Beispiel inTouch“, „The Bed – das Bett“, und „ComTouch“ heran, fasse deren Inhalt zusammen und wage im Anschluss einen Ausblick in die Zukunft dieser Forschungsrichtung.

2. Umarmung in die Ferne: eine neue Art der Kommunikation

2.1 Idee der entfernten Umarmung

Diese Arbeit stellt ein Produkt vor, welches sich „The Hug“ [1] nennt. Eine Art Kissen mit Armen soll dem menschlichen Bedürfnis nach körperlicher Nähe genüge tun (vgl. Fig. 1). Gerade ältere, einsame Menschen sollen durch diese Erfindung die Möglichkeit haben, einen Schritt aus Ihrer Einsamkeit zu tun, und - wenn auch auf eine abstrakte Weise - körperliche Nähe in Form einer Umarmung zu geliebten Verwandten aufzubauen, bzw. über die Ferne hinweg aufrechtzuerhalten.



Fig. 1. Prototyp „The Hug“[1]

Dieses Produktkonzept erlaubt es, körperliche Grüße zu versenden. Wird z.B. eine Umarmung von der Großmutter an die Enkelin geschickt, klingelt dieser „Umarmungsapparat“ des Kindes und leuchtet auf. Die Enkelin hat nun die Möglichkeit, die Umarmung durch drücken der linken Pfote anzunehmen. Wie über ein Telefon können die zwei nun miteinander sprechen, aber den Hug zusätzlich auch streicheln und drücken. Dieser übersetzt diese Gesten in Vibration und Wärmeentwicklung. Es wurde viel Aufwand in das äußerliche Design des „Hug’s“ gesteckt, und sogar ein Sounddesigner für die verfügbaren Geräusche bemüht.

Auf diese Weise agiert das Gerät als Stellvertreter für eine nicht anwesende Person. Eine Art Proxy also, der Gefühle annimmt, in technische Rohdaten umwandelt, zu seinem Äquivalent schickt, dort wieder in vergleichbare, körperliche Äußerungen zurückverwandelt und der adressierten Person mitteilt.

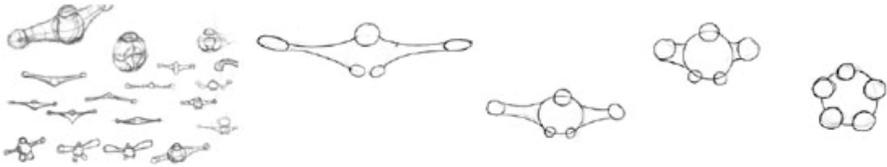


Fig. 2. Entwicklungsstudien zu Form und Design des Hugs [1]

2.2 Funktionen und Handhabung

Die technischen Hauptfunktionen des Hugs sind interpersonelle Kommunikation mittels kabelloser Telefonie und Drucksensoren. Eine Verbindung wird aufgebaut, indem man die linke Pfote drückt und den Namen des Empfängers in das Mikrofon spricht (vgl. Fig. 3). Man beendet die Verbindung, indem man die rechte Pfote drückt und somit auflegt. Es werden bei jedem Gespräch die Daten des Anrufers in einem so genannten „Hug-Netzwerk“ gespeichert. Dieses besteht aus nur wenigen, sehr vertrauten Personen des engsten Bekanntenkreises.

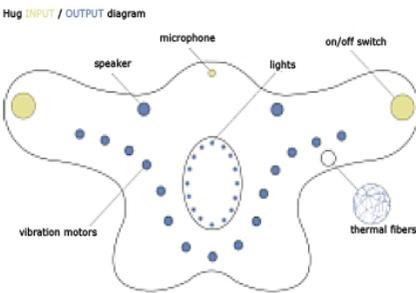


Fig. 3. Technische Details [1]

Bis jetzt wird der Hug noch nicht hergestellt. Diese Arbeit basiert letztlich nur auf der Vorstellung der Wichtigkeit gefühlter Technologie um soziale und emotionale Kommunikation zu ermöglichen. Der Hug wurde speziell für die grundlegenden Bedürfnisse nach Nähe zu geliebten Personen entwickelt. Er soll Technologie zu den Menschen bringen, die sonst nicht viel mit diesem Fachbereich zu tun haben und gleichzeitig bestehende Ängste und falsche Erwartungen gegenüber dem Medium aufklären. Doch am Wichtigsten ist die soziale und emotionale Unterstützung, die durch diese Erfindung alten Menschen zuteil werden könnte (vgl. Fig 4).

Es stellt sich die Frage, ob die vom Hug repräsentierten Kanäle als „Menschenfreundlich“ angenommen werden oder eher befremdlich wirken? Durch die Verschmelzung mehrerer Kommunikationskanäle (taktile, visuelle, auditive, thermale) erscheint die Idee zumindest schon innovativer als Kommunikationslösungen, die sich bislang nur auf einen Kanal konzentrierten. Dadurch, dass man die einzelnen Kanäle auch unabhängig von einander bedienen und verschicken kann (z.B. kann eine einfache Umarmung als allein stehender Gruß versendet werden, ohne dass man ein Telefonat führen muss), kommen auch Menschen als Zielgruppe in Frage, die die gängigen Kommunikationstechniken auf Grund von Unvermögen nicht gerne oder gar nicht benutzen.

Es stellt sich die Frage, ob die vom Hug repräsentierten Kanäle als „Menschenfreundlich“ angenommen werden oder eher befremdlich wirken? Durch die Verschmelzung mehrerer Kommunikationskanäle (taktile, visuelle, auditive, thermale) erscheint die Idee zumindest schon innovativer als Kommunikationslösungen, die sich bislang nur auf einen Kanal konzentrierten. Dadurch, dass man die einzelnen Kanäle auch unabhängig von einander bedienen und verschicken kann (z.B. kann eine einfache Umarmung als allein stehender Gruß versendet werden, ohne dass man ein Telefonat führen muss), kommen auch Menschen als Zielgruppe in Frage, die die gängigen Kommunikationstechniken auf Grund von Unvermögen nicht gerne oder gar nicht benutzen.



Fig. 4. körperliche Nähe durch Verbindung über das Telefonnetz? [1]

3. Das Bett als ein Medium für intime Kommunikation

3.1 Hintergrund

„The Bed“, das Bett stellt ein weiteres Produkt dar, welches den Menschen Nähe zu bestimmten anderen Personen vermitteln soll [2]. Das Ganze kann wieder nur auf abstrakte Weise funktionieren und das Medium über welches die Idee diesmal realisiert werden soll, ist das Bett. Der Verfasser erwähnt die zahlreichen emotionalen Verbindungen, die jeder Mensch in seinem Leben mit den einzelnen Komponenten seines Bettes assoziiert. Allein das Kissen diene neben einer Kopfunterlage auch oftmals als Trostpender in traurigen Zeiten, als Sandsack während eines Wutanfalls, oder als Partnerersatz zum Kuschneln. Die zweite Eigenart, die sich der Entwickler zunutze machen will, ist die Bedeutung eines Vorhangs, der sich durch Bewegung im Raum meist selbst mitbewegt und somit die Anwesenheit eines anderen Individuums repräsentiert. Neben dem bekannten Anbringungsplatz vor dem Fenster wird er auch oftmals vor das Bett gehängt.

Obwohl sie keine physischen Objekte sind, will der Autor auch Schatten als Indikator für ungesehene, anwesende Körper in das Kommunikationsmittel Bett einbringen. Jeder Mensch fülle diese Formen und Bewegungen mit eigenen Phantasien aus. Wie aber kann man sich diese vertrauten Komponenten dienlich machen und sie als umgewandelte Daten einer entfernten Person übermitteln?

3.2 Umsetzung

Mit den vorangehenden Erläuterungen soll ein funktionierender Prototyp der Bett-schnittstelle präsentiert werden, der die Lücke zwischen körperlicher und abstrakter, virtueller Anwesenheit schließt.

Die Implementierung sieht zwei Betten vor, mit jeweils zwei Kissen und einem Vorhang. Die zwei Betten sind über TCP/IP miteinander verbunden. Jedes dieser Kissen erfüllt seinen eigenen Zweck. Das eine dient klassisch für den Kopf, das andere kann an den Körper gepresst werden. Neben kleinen, installierten Miniboxen zur Audioausgabe befinden sich zwei Sensoren im Kopfkissen, die für die Eingabe zuständig sind (vgl. Fig. 5). Der eine Sensor ist ein Mikrofon, der andere misst den genauen Standpunkt des Körpers im Bett. Das zweite Kissen für den Körper ist ein reines Ausgabegerät um als physischer Stellvertreter des entfernten Individuums zu dienen. Eine

elektrische Heizdecke strahlt Wärme aus und ein großer Subwoofer kann starke Schwingungen erzeugen.

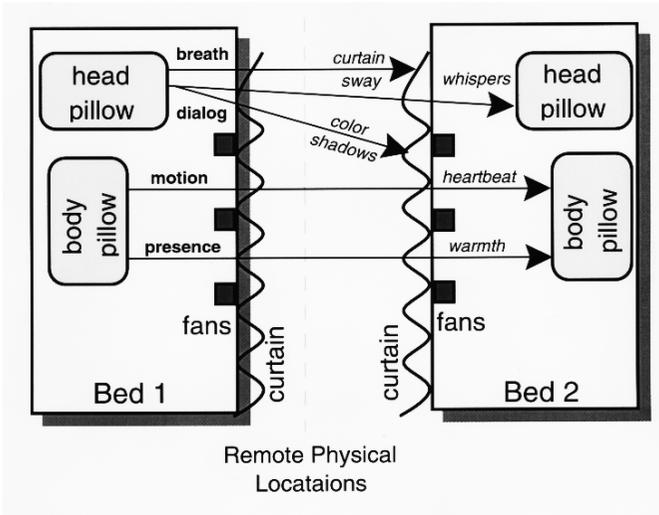


Fig. 5. entfernte Kommunikation im Bett [2]

Schließlich hängt noch ein halbtransparenter Vorhang gleich neben dem Bett, auf den Verschiedenes projiziert werden kann. An dessen unterem Ende sind eine ganze Reihe von computergesteuerten Ventilatoren angebracht die auf Kommando Luft gegen den Vorhang blasen. Durch einen bestimmten Aufbau von Plexiglasstrahlern kann zusätzlich noch ein beliebiges Farbenspiel erzeugt werden, welches sich auf dem Vorhang als Schatten abzeichnet.

Die Bewegungen eines Bettbenutzers werden nun in diese abstrakten, umgebungsabhängigen Reaktionen übersetzt (vgl. Fig. 6). Da die Betten bidirektional miteinander verbunden sind, kann zeitgleich kommuniziert werden. Halten beide Teilnehmer in ihren Betten das Kissen an den Körper gedrückt, wird die Anwesenheit des anderen durch eine konstante Wärme und ein langsames, deutlich pulsierendes Signal repräsentiert. Auf diese Art wird die physische Distanz der beiden Agierenden überbrückt, wobei der eine vom anderen sowohl den Puls als auch seine Körperwärme spüren kann. Sobald sich eine Person in ihrem Bett unruhig hin und her wälzt, wird dies dem Partner durch einen erhöhten Pulsschlag übermittelt.

Wenn etwas gesprochen wird, analysiert ein Computer die Qualität der Audiodaten in Bezug auf die Amplitude (Lautstärke, Sanftheit) und überträgt diese „Qualitäten“ (nicht die Audiodaten selbst) zu dem anderen Bett. Dort werden die angekommenen Daten ausgewertet und als gefärbte Schatten auf die Vorhänge projiziert. Dies ermöglicht eine Farbkomposition durch Sprache und abstrahiert somit die Möglichkeiten der Kommunikation hochgradig.

Die Audioanalyse unterscheidet zusätzlich zwischen gesprochener Sprache und „gehauchten“ Lauten, wie Atmung oder Pusten. Sollte ein Teilnehmer letzteres anwen-



Fig. 6. Das Bett im Einsatz [2]

den, wird beim entfernten Bettpartner entsprechend der Vorhang durch die Ventilatoren bewegt, so als hätte der Atem die Entfernung überbrückt, um dem Empfänger die eigene Nähe künstlich darzustellen.

4. Entferntes Fühlen und Tasten

4.1 Vorbemerkung

In der nun folgenden Arbeit wird die Berührung ebenso als fundamentales Mittel zur Festigung der Kommunikation, ja gar als nicht wegzudenkendes Kommunikationsmittel selbst angesehen. „...physical contact can convey a vitality and immediacy at times more powerful than language.“ (Brave, Ishii, Dahley, S. 4). Auch hier wird fest davon ausgegangen, dass eine Erweiterung der Kommunikation durch entfernte Berührung ein zukunftsreicherer Zweig sei und der Entwicklung und Forschung bedarf. Aus diesem Grund stellen die Autoren ein Gerät namens inTouch vor, das haptische, interpersonelle Kommunikation ermöglichen soll, welche räumliche Distanz überbrückt.

4.2 Design von inTouch

Das Modell von inTouch besteht aus zwei bidirektional verbundenen, handtellergroßen Geräten mit jeweils drei zylindrischen Rollen (vgl. Fig. 7). Jeder Rolle ist mit seinem Äquivalent auf der anderen Seite verkoppelt. Durch eine Force-Feedback Technologie und Positionssensoren wird auf einer Seite festgestellt, in welchem Zustand sich die betroffene Walze befindet, um dann mittels hoch präzisen Motorenden Gegenpart anzupassen. Dadurch können zwei interagierende Individuen fühlen, wie der jeweilige Antagonist eine dieser drei Rollen, beeinflusst (vgl. Fig. 9).

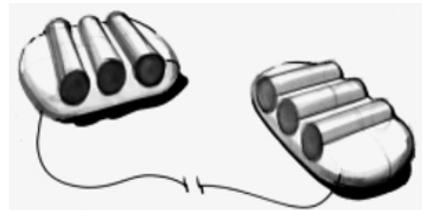


Fig. 7. erste Vorstellung von „inTouch“ [3]

4.3 Umsetzung und Erweiterungen

Es wurden nacheinander drei Prototypen realisiert. Der erste (inTouch-0) setzte den Walzenabgleich über elastische Antriebswellen um (vgl. Fig. 8). Die Klasse der Studenten, in der dieser erste Typ vorgestellt wurde, beschrieb den Umgang mit ihm als spielerisch. Es wurde aber auch sogleich die fehlende Möglichkeit zu weiterführendem Informationsaustausch bemängelt.

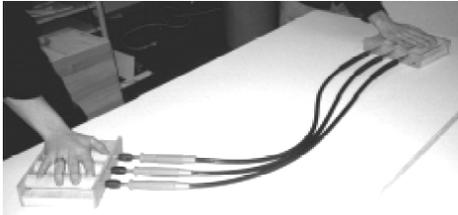


Fig. 8. mechanische Attrappe [3]



Fig. 9. Prototyp "inTouch" [3]

Im zweiten Prototyp, dem inTouch-1 bediente man sich schon einer ausgefeilteren Technik. Eine Kombination aus Hewlett Packard's optischen Positionsüberwacher, der den physischen Stand der Walzen beobachtet und misst, Maxon DC Motoren, die die Synchronisation durchführen und einem Kontrollalgorithmus der auf einem Host PC läuft, und die Unterschiede der verbundenen Rollen überwacht und gegebenenfalls ausgleicht. (vgl. Fig. 10)

Der zweite Prototyp wurde vom Entwicklerteam sowohl auf einem Sponsorentreffen ausgestellt, als auch intern getestet. Mehr als 500 Menschen probierten dieses Modell aus und die überwiegende Mehrzahl der Reaktionen war äußerst positiv. Viele glaubten sogar, dass inTouch dazu beitragen könne, den emotionalen Zustand entfernter Personen zu errahnen. Dies wurde jedoch (noch) nicht weiter getestet.

Die dritte Implementierung, der inTouch-2 unterscheidet sich nur insofern von inTouch-1, als das die Positionen der Rollen auf entfernte PCs übertragen werden, und somit eine Verbindung des inTouch-Systems über das Internet möglich wurde (vgl. Fig 11). Außerdem wurden kleine Veränderungen des Kontrollalgorithmus' angepasst.

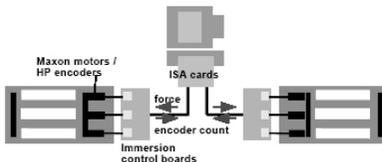


Fig. 10. inTouch-1 Systemarchitektur [3]

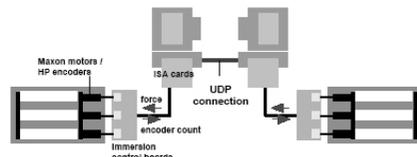


Fig. 11. inTouch-2 Systemarchitektur [3]

5. Telephonie durch Berührung anreichern

5.1 Erläuterung

Während in den vorangegangenen Arbeiten immer eigenständige Apparate entwickelt wurden, welche neue Kommunikationskanäle nutzen, wird in folgender Arbeit eine Idee vorgestellt, welche bestehende Kommunikationsmittel über den auditiven Kanal unterstützen und somit zu verbesserter interpersoneller Verständigung führen soll [4]. Es wurde nicht nur der Prototyp eines Geräts entwickelt, sondern auch dessen Wir-

kung auf das Kommunikationsverhalten von 24 Probanden untersucht. Aufgrund der Ausbreitungs natur von Audio sei der Mensch vielerlei Sinnesüberreizungen ausgesetzt, da von überall auditive Information auf ihn eindringt und dadurch Informationsverlust auftritt. Um dies zu umgehen, entstand die Idee von ComTouch, einem Zusatzgerät, welches die Sinnesmodalität der Berührung zur Unterstützung auditiver Kommunikation einsetzen soll (vgl. Fig. 12).

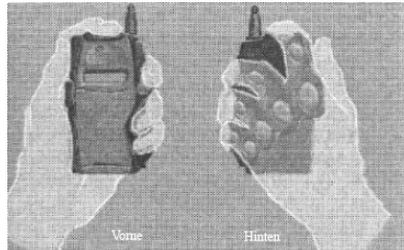


Fig. 12. konzeptioneller Zeichnungsentwurf
- So könnte ComTouch in Zukunft aussehen [4]

5.2 Versuchsaufbau und Schlussfolgerung

Das grundlegende hardwarespezifische Konzept basiert auf einer Art Manschette, die über die Rückseite eines Handys gezogen werden kann. Theoretisch soll nun ein Benutzer durch Fingerdruck eine Vibration erzeugen können, die das verbundene Gerät ausgibt. Durch eine bidirektionale Verbindung soll es möglich sein, Vibrationen zu schicken und auch zu empfangen.

Um die Effektivität und den Nutzen von taktiler Kommunikation beurteilen zu können, werden die Kommunikationsarten Seh- und Hörbehinderter beobachtet. Es werden die Techniken des Fingerspellings (Buchstabieren über Handkontakt), des Tado-ma (der Daumen wird auf die Lippen gelegt und durch Lippenbewegung und Atemstöße der Inhalt der Wörter erkannt) und des Braille (auf einer ebenen Unterlage werden hervortretende Punkte als Buchstaben abgetastet) erwähnt, um die enge Verbindung zwischen Verständigung und Berührung aufzuzeigen. Dadurch soll auch gezeigt werden, dass sich Berührung theoretisch auch als primäres Kommunikationsmedium eignet (vgl. Fig. 13 und Fig. 14).

Es galt, eine Methode zu entwickeln, die ein entferntes Berühren ermöglicht. Als Mittel zur Abstraktion der Berührung wurde ganz bewusst die Vibration gewählt, um einen Händedruck abstrahiert darzustellen. Außerdem wurde diese schon in vielen

Im darauf folgenden Versuchsaufbau wurden je drei dieser Vibrationserzeuger auf zwei verbundene Platten angebracht, welche als Schnittstelle für beide Anwender diente. Nur ein Finger hatte die Möglichkeit zu kommunizieren, was zur Vereinfachung durch drei farbige Markierungen verdeutlicht wurde (vgl. Fig. 16).

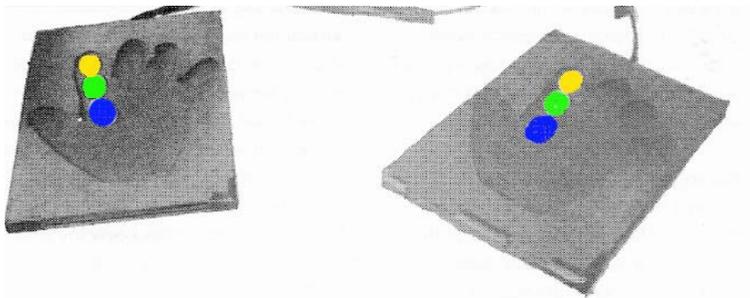


Fig. 16. vorausgehender Versuchsaufbau zu ComTouch [4]

Der gelbe, obere Teil diente als Sender. Je nach Druck konnte man hier unterschiedlich starke Vibrationen erzeugen. Der grüne, mittlere Teil war die Vorrichtung zur Bestätigung des gerade gesendeten Signals, also ein Feedback darüber, wie stark der eigene Druck beim verbundenen Partner ankommt. Der blaue, untere Part diente als Empfang für eingehende Signale.

Den 24 Probanden wurden zusätzlich zu diesem Versuchsaufbau jeweils ein Kopfhörer und ein Mikrophon für den auditiven Kontakt zur Verfügung gestellt. Nun sollten sie verschiedene Aufgaben lösen, damit mehr Wissen über die Verbindung von taktilem und auditivem Kommunikation gesammelt werden konnte.

Beobachtungen der Testpersonen und ihres Verhaltens ergaben interessante Ergebnisse, wie das intuitiv schnelle Verständnis der Handhabung, oder die Hervorhebung mancher Wörter durch Drücken des gelben Sendeknopfes (vgl. Fig. 17).

Kurz vor der Wortübergabe an den Gesprächspartner, wurde von diesem fast immer ein taktiles Signal registriert, welches als Zeichen für eine Wortmeldung interpretiert wurde. Eine Art von Spiel entwickelten die meisten der Probanden selbstständig, indem sie sich gegenseitig rhythmische Vibrationsabfolgen vorgaben und diese dann nachahmten. 83% der 24 Personen gaben an, mindestens eine der soeben beschriebenen Kommunikationsmethoden benutzt zu haben.

Die freiwilligen Teilnehmer waren sich untereinander vertraut und somit wurde von den Autoren keine Aussage über die taktile Kommunikation von Fremden getroffen. Die Ergebnisse jedoch, zu denen sie bei diesen Versuchen gelangten, führte sie zu der Schlussfolgerung, dass es einerseits noch viel zu verbessern gäbe (sei es die ergonomische Form, die Anzahl der Kanäle, die Genauigkeit der Vibrationsstärken oder die Lautstärke der Vibrationen selbst), andererseits aber feststehe, dass es eine enge Be-

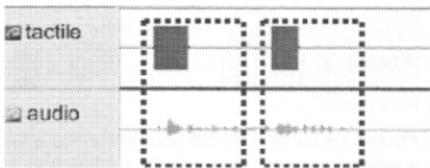


Fig. 17. Sätze, Wörter und sogar Silben wurden durch die Vibration verstärkt [4]

ziehung zwischen taktilem und auditivem Kommunikation.

ziehung zwischen dem auditiven und dem taktilen Kanal gäbe und die Zusammenführung dieser beiden nur zum Wohle der kommunikativen Gesellschaft wäre. Ihnen zufolge wäre eine Zukunft vorstellbar, in der Handys mit einer Art ComTouch ausgestattet werden, um sogar Sinnesgeschädigten Menschen die entfernte Kommunikation zu erleichtern.

6. Diskussion und Vergleich

Durch die Untersuchung vier verschiedener Entwicklungen und deren Umsetzung wird klar, dass diese tendenziell das gleiche Ziel verfolgen: die bestehenden Techniken zur entfernten Kommunikation sollen wie im Falle von ComTouch ausgereizt und erweitert, oder wie in den anderen Beispielen als Grundlage benutzt werden, um gänzlich neue Arten zu entwickeln.

Während mit dem „Bed“ und dem „Hug“ ein konkreter Produktvorschlag gemacht wurde, speziell entwickelt für bestimmte Situationen in denen das Leben leichter gemacht werden soll, so haben „ComTouch“ und „InTouch“ das Ziel, experimentell die menschlichen Sinneskanäle auszunutzen, um interpersonelle, entfernte Kommunikation auf eine neue Ebene zu führen.

Der Hug ist als Gerät insofern vorstellbar und sinnvoll, als dass er Menschen hilft, Beziehungen auf einer Ebene zu pflegen, die durch ein Telefon allein nicht realisierbar wären. Intuitiv kann man sich das Konzept des Einsamkeitsmindernden Gerätes als sehr erfolgreich vorstellen, da es ein menschliches Grundbedürfnis ist, die Nähe zu geliebten Personen zu suchen. Ob aber eine Art Kuschelroboter als Stellvertreter für Zärtlichkeiten von alten Menschen angenommen würde, ist jedoch fraglich. Neugierde weckt das Konzept allemal, und dazu Lust auf mehr dieser Vorschläge und Ideen.

Das „Bed“ erfüllt diese intuitiven Vorstellungen vielleicht sogar noch mehr, da die Entwickler kein neues, vielleicht abschreckendes Gerät (wie dem Hug) einführen, sondern ein bereits bestehendes, allseits vertrautes und täglich genutztes Objekt erweitern, um das Gefühl von Nähe zu übertragen. Dieses Projekt steckt, genau wie der „Hug“ auch, voller innovativer, interessanter Ideen. Durch die Darstellung von Nähe durch die Abstraktion des Pulsschlages wird die ganze Konstruktion, die ja doch auf Computer basiert für den Nutzer schnell sehr menschlich. Einzig die Spielerei mit den Schatten mag nicht so ganz einleuchten, stehen diese doch oftmals für alles andere als Beruhigung. So kann es gut sein, dass der gewünschte Effekt verloren geht, wenn er nicht sogar ins Gegenteil umgewandelt wird.

Diese beiden Kommunikationsgeräte sind weiterhin sehr gut vorstellbar, wenn zwei Menschen eine Fernbeziehung führen. Der schleichenden Verfremdung durch Nicht-anwesenheit kann mit solchen Mitteln vermutlich durchaus vorgebeugt werden, und wer einmal eine geliebte Person in weiter Ferne hatte, der kann auch nicht bestreiten, dass sich in einem selbst die Nachfrage nach so einer Erleichterung des zwischenmenschlichen Kontaktes regt.

Auch für die Überbrückung von Distanzen geeignet ist natürlich inTouch. Auch hier wird Berührung und somit Nähe übertragen. Allerdings wurde hier nicht erforscht, warum das den Menschen gefallen könnte, sondern wie man die Berührung an sich am Besten übertragen kann. Die Endgeräte bestehend aus den drei Rollen mit denen der Benutzer in Kontakt kommt, erscheinen für die Übertragung von Berührung fast zu grob. Doch auf dem Ansatz aufbauend lassen sich schnell neue Ideen und Weiterentwicklungen finden.

Anhand der Ergebnisse der Versuche von ComTouch wird schnell klar, dass sich die Forschung in diesem Bereich noch in einem ziemlich frühen Stadium befindet. So kann ein kritischer Leser anführen, dass es nur 24 Testpersonen waren, die sich auch noch alle gekannt haben. Die Ergebnisse jedoch sind beeindruckend und fast wünscht man es sich, schon heute das ins Handy gesprochene Wort, durch einen festen Fingerdruck zu untermalen.

Die ersten Ansätze dazu sind da, und wer ein wenig Phantasie hat, kann sich vorstellen, wie es weitergehen kann. Wenn sich die richtigen Ideen durchsetzen, wird es vielleicht schon in naher Zukunft möglich sein, auf mehreren Kanälen als nur dem sprachlichen miteinander zu kommunizieren, sich Gesten zu schicken, Berührung zu übertragen und einfach näher bei jemand zu sein, der in Wirklichkeit sehr weit entfernt ist.

Literatur

- [1] DiSalvo, C., Gemperle, F., Forlizzi, J., und Yonkers, W. "The Hug: a New Form for Communication", veröffentlicht in "DUX 2003 case studies", 2003 (<http://www.aiga.org/resources/content/9/7/8/documents/gemperle.pdf>)
- [2] Dodge, C. "The Bed: A Medium for Intimate Communication", Cambridge, 1997, (<http://sigchi.org/chi97/proceedings/short-talk/cd.htm>),
- [3] Brave, S., Dahley, A., Ishii, H. "Tangible Interfaces for Remote Collaboration and Communication", veröffentlicht in den "Proceedings of CSCW", November 1998, (http://tangible.media.mit.edu/papers/inTouch_PSyBench_CSCW98/inTouch_PSyBench_CSCW98.pdf)
- [4] Chang, A., Gunther, E., Ishii, H., Jacob, R., O'Modhrain, S. "ComTouch: Design of a Vibrotactile Communication Device", London, 2002, (http://www.medialabeurope.org/palpable/comtouch_dis.PDF)
- [5] LEO - Link Everything Online, Dictionary, <http://dict.leo.org/>
- [6] Heinz Pürer – Einführung in die Kommunikationswissenschaft Teil 1
- [7] http://www.diversplatform.net/eng/conversion/pressure/e_bar_psi.pdf

Benutzerschnittstellen für Sprache und Stimme

Raphael Wimmer

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
wimmer@informatik.uni-muenchen.de

Zusammenfassung Sprach-Benutzerschnittstellen ermöglichen es dem Menschen, in natürlicher Sprache mit dem Computer zu interagieren. Für mobile Endgeräte wie Mobiltelefone oder Persönliche Digitale Assistenten (PDAs) bietet Sprache eine kostengünstige und anwenderfreundliche Benutzerschnittstelle. Die grundlegenden Entwurfsprinzipien unterscheiden sich aber erheblich von denen für grafische Benutzerschnittstellen. Diese Arbeit stellt vier interessante Forschungsarbeiten über auditive Benutzerschnittstellen vor. Diese behandeln Entwurfskriterien für Sprachdialogsysteme, Techniken für Sprach-Benutzerschnittstellen in mobilen Endgeräten und alternative Eingabemöglichkeiten durch nonverbale Laute. Eine Zusammenstellung von grundlegenden Entwurfsprinzipien für Sprach-Benutzerschnittstellen ergänzt die Fallstudien. Zusätzlich zu den akademischen Arbeiten liefert diese Arbeit eine Marktübersicht über kommerzielle und freie Sprachsoftware und eine kurze Vorstellung von VoiceXML, der Standard-Beschreibungssprache für Sprachdialogsysteme.

1 Einleitung

Noch vor der Erfindung der Schrift war Sprache die erste Möglichkeit des Menschen, mit Anderen zu kommunizieren. Bei der Interaktion mit dem Computer führt die gesprochene Sprache allerdings ein Schattendasein. Grafische Benutzerschnittstellen (Graphical User Interfaces, GUIs) sind dagegen allgegenwärtig. Sie haben kein Gegenstück in der realen Welt (außer andere Maschinen). Insofern war es für den Menschen nur natürlich, seine Arbeitsweise an die der Anwendung anzupassen. Bei Sprach-Benutzerschnittstellen (Speech User Interfaces, SUIs) stehen die Entwickler aber vor der Herausforderung, deren Struktur an die der menschlichen Sprache anzupassen. Denn genauso wie man von einem anderen Menschen erwartet, dass er einfache Befehle und Aussagen versteht, erwartet der Benutzer implizit, dass auch ein Sprachdialogsystem ihn versteht, wenn er sagt "Nein, nicht diese Email öffnen, die andere von, äh, von dem Vertreter, der gestern da war". Diese Erwartungshaltung stellt die Entwickler von Sprach-Benutzerschnittstellen vor die Herausforderung, ein System, das bei weitem nicht die geistige Leistungsfähigkeit eines Menschen hat, so zu gestalten, dass es als Dialogpartner agieren kann. Aus diesem Grund sind Sprach-Benutzerschnittstellen für die Usability(Benutzerfreundlichkeits)-Forschung ein

fruchtbares Gebiet. Die vier hier vorgestellten Arbeiten zeigen ein breites Spektrum an Ideen, um natürlichsprachige Kommunikation mit dem Computer benutzerfreundlich zu gestalten. Vor allem für mobile Endgeräte bieten sich Sprach-Benutzerschnittstellen an, da sie keinen Platz für Bildschirm und Tastatur benötigen und auch parallel zu anderen Tätigkeiten benutzt werden können.

Diese Arbeit geht vor allem auf die folgenden vier wissenschaftlichen Veröffentlichungen ein. Yankelovich, Levow und Marx berichten in *"Designing SpeechActs: Issues in Speech User Interfaces"* [3] von ihren Erfahrungen bei Design, Implementierung und Evaluation einer Sprach-Benutzerschnittstelle für E-Mail, Kalender, Börsenkurse und Wetterbericht. In *"Nomadic Radio: Speech & Audio Interaction for Contextual Messaging in Nomadic Environments"* [4] stellen Nitin Sawhney und Chris Schmandt vom MIT Media Lab ihre Designprinzipien für eine mobile sprachgesteuerte Kommunikationsplattform dar und gehen besonders auf kontextbasierte Benachrichtigungen und unterschwellige Informationsdarstellung ein. *Voice User Interface Principles for a Conversational Agent* [1] von Ross, Brownholtz und Armes stellt schließlich allgemeine Design-Regeln für Sprachdialogsysteme auf. Takeo Igarashi und John F. Hughes beschreiben in *"Voice as Sound: Using Non-verbal Voice Input for Interactive Control"* [2], wie man mit Lauten statt durch Sprache Geräte steuern kann.

Ein ausführlicher Überblicksartikel zum Thema Sprachdialogsysteme ist *"Spoken Dialogue Technology: Enabling the Conversational User Interface"* [6] von Michael F. McTear.

2 Sprach-Benutzerschnittstellen - Begriffserklärung und Eingrenzung

2.1 Sprach-Benutzerschnittstellen

Sprach-Benutzerschnittstellen (Voice User Interfaces / Speech User Interfaces, SUI) stellen dem Benutzer Methoden zur Verfügung, um mit dem Computer über gesprochene Sprache zu interagieren.

2.2 Sprachdialogsysteme

Sprachdialogsysteme (Spoken Dialogue Systems) sind Computersysteme, mit denen Menschen interagieren können und in denen natürliche Sprache eine wichtige Rolle für die Kommunikation spielt [6]. Sie verwenden eine Sprach-Benutzerschnittstelle, um den Zugriff auf eine Anwendung oder Datenquelle zu realisieren.

Ein- und Ausgabe können sich unterschiedlich stark an die natürliche Sprache anlehnen. Aus praktischen Gründen werden zum Beispiel oft nur "Ja", "Nein" und die Zahlen von Null bis Neun zur Eingabe verwendet. Die Ausgabe kann auch lediglich aus einer Aktion bestehen ("Computer, medizinisch-holographisches Notfallprogramm aktivieren").

Ein Sprachdialogsystem besteht üblicherweise aus sechs Komponenten (Abbildung 1).

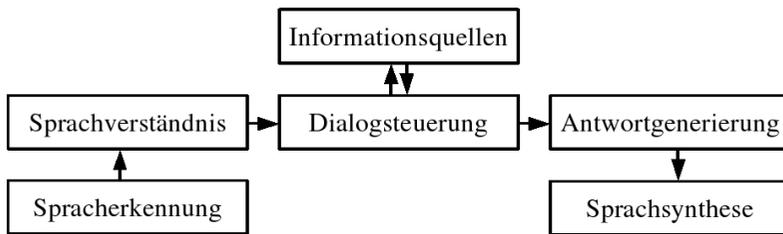


Abbildung 1. Üblicher Aufbau eines Sprachdialogsystems

Spracherkennung. Eine Spracherkennungssoftware (Speech Recognition Engine) wandelt gesprochene Sprache in Text um (z.B. "Überweise 500 Euro an..."). Dafür werden meist Hidden Markov Models (HMM) [10] oder Dynamic Time Warping (DTW) [9] eingesetzt. Spracherkennung ist immer noch das Sorgenkind der Entwickler, da die durchschnittliche Erkennungsrate heutiger Implementierungen unter realistischen Bedingungen bei 80%-90% liegt. Sind laute Hintergrundgeräusche vorhanden - bei Mobiltelefongesprächen keine Seltenheit - sinkt die Erkennungsrate sogar bis auf 50%.

Sprachverständnis. Nachdem die Äußerung des Benutzers als Text vorliegt, kann ein Syntaxanalysator (Parser) diese auf bedeutungsvolle Inhalte analysieren. Der Algorithmus muss feststellen, welche Daten bzw. Befehle der Text enthält (z.B. Aktion="Überweisung", Betrag="500,00", ...)

Dialogsteuerung. Die Dialogsteuerung erhält die aufbereiteten Daten vom Syntaxanalysator und verarbeitet diese. Dabei können gespeicherte Daten verändert werden (z.B. Kontodaten) und Ausgaben vorbereitet werden (z.B. eine Bestätigungsmeldung).

Informationsquellen. Die Dialogsteuerung hat Lese- und Schreibzugriff auf eine Datenbank. Diese kann Informationen liefern (z.B. Benutzername="Herr Meyer") und auch Datenänderungen entgegennehmen (z.B. Benutzer[4534].Konto[00].Betrag = "1242,00")

Antwortgenerierung. Nachdem die Dialogsteuerung eine passende Antwort auf die Anfrage ermittelt hat (z.B. Status="ÜberweisungOK"), wird ein Antworttext erzeugt (z.B. "Vielen Dank für Ihre Überweisung. Was kann ich sonst noch für Sie tun?").

Sprachsynthese. Der Antworttext wird von einer "Text-To-Speech(TTS)-Engine" in eine Sprachausgabe umgewandelt. Entweder werden dazu einzelne Phoneme (atomare Bestandteile gesprochener Sprache) zu Wörtern zusammengesetzt, oder vorgefertigte Sprachbausteine aneinandergehängt. Alternativ kann auch ein vorher aufgenommener Text ausgegeben werden.

2.3 Beschränkung auf Usability-Forschung

Sprachdialogsysteme sind ein komplexes Thema mit vielen sehr unterschiedlichen Problemstellungen. Während vor allem die Spracherkennung und Sprachsynthese in den letzten zehn Jahren schon eine erhebliche Entwicklung durchgemacht haben, erhält inzwischen auch ein anderer Aspekt von Sprachdialogsystemen verstärkte Beachtung: Die Benutzerfreundlichkeit (Usability) der Systeme. Sprach-Benutzerschnittstellen gehorchen gänzlich anderen Prinzipien als grafische Benutzerschnittstellen. Zum Beispiel sieht der Benutzer während einer Sitzung mit einem Sprachdialogsystem nicht, welche Felder eines Formulars er schon ausgefüllt hat. Er muss sich also ein mentales Modell des Formulars bilden. Diese und weitere Herausforderungen im Design von Sprach-Benutzerschnittstellen behandeln die Arbeiten, die im Folgenden vorgestellt werden.

3 SpeechActs: Eine Sprach-Benutzerschnittstelle für Messaging und Kalender

3.1 Was ist SpeechActs?

SpeechActs war 1995 ein Forschungsprototyp für Sprachdialogsysteme bei Sun Microsystems. Das System stellt dem Benutzer über ein Sprachdialogsystem den Zugriff auf seine E-Mails, seinen Kalender, Börsen- und Wechselkurse und den Wetterbericht zur Verfügung. Die Sprachausgabe arbeitet mit einer synthetisierten Computerstimme.

Beispielablauf:

SpeechActs: Sie haben 14 neue und 47 alte Nachrichten. Die erste neue Nachricht ist von Eric Baatz, Betreff "Bericht fertig".

Benutzer: Will ich hören.

SpeechActs: "Der erste Entwurf wartet auf deine Kommentare. Eric"

Benutzer: Antwort inklusive Originalnachricht.

SpeechActs: Bereite Nachricht an Eric Baatz vor. Bitte sprechen Sie nach dem Piepton. Wenn Sie fertig sind machen Sie eine Pause von mehreren Sekunden

Benutzer: Eric, ich schick's dir heute nachmittag.

SpeechActs: Sagen Sie Abbrechen, Senden oder Vorschau.

Benutzer: Senden.

SpeechActs: Nachricht gesendet. Was nun?

Benutzer: Nächste Nachricht.

SpeechActs: Nachricht zwei von Stuart Adams, Betreff: "Können wir uns um drei treffen?".

Benutzer: Wechsle zum Kalender.

...

3.2 Evaluation

Im Laufe der Entwicklung wurden iterative Usability-Tests durchgeführt. Als Versuchsaufbau diente ein Raum mit Telefon, das mit dem SpeechActs-System verbunden war. Den Teilnehmern wurde aufgetragen, innerhalb von 20 Minuten 22 Aufgaben über das SpeechActs-System zu erledigen. Diese umfassten das Lesen und Beantworten von Mail, das Aufrufen von Kalendereinträgen, Börsenkursen und Wetterberichten. Ein Problem, das dabei festgestellt wurde, war die nicht zufriedenstellende Qualität der Spracherkennung. Auch waren die Benutzer mit dem Mail-System von Sun vertraut, das seine Eingaben über die Wähltastatur entgegennahm. Oft wurde deshalb von den Teilnehmern angemerkt, dass eine kurze Tastensequenz viel schneller gewesen wäre als die Sprach-Kommandos.

3.3 Herausforderungen

Im Laufe der Entwicklung und Evaluation wurden einige Prinzipien klar, die beachtet werden müssen, um ein benutzerfreundliches Sprachdialogsystem zu bauen.

Konversation imitieren Um die Benutzung einfach zu gestalten, muss sich der Computer an das menschliche Konversationsverhalten anpassen, anstatt ihm die traditionelle Arbeitsweise einer grafischen Anwendung aufzuzwingen.

So sollte es ein Sprachdialogsystem vermeiden, den Benutzer explizit zu einer Eingabe aufzufordern, sondern ihm die freie Wahl über seine Aktionen lassen. Die führt allerdings auch zu einer hohen Zahl unterschiedlicher Situationen und Verzweigungen in der Gesprächsführung, die von der Software behandelt werden müssen.

Gesprächstempo Besonders wichtig für regelmäßige Benutzer ist die Geschwindigkeit, mit der er seine Aufgaben über eine Sprachschnittstelle erledigen kann. Ein Benutzer, der zum ersten Mal mit einem Dialogsystem interagiert, ist zweifellos dankbar für ausführliche Fragen und Antworten des Systems. Bei regelmäßiger und intensiver Nutzung eines Dialogsystems werden häufig vorkommende Ausgaben des Dialogsystems wie "Sie haben eine neue Nachricht erhalten" dem Benutzer schnell lästig. Aus diesem Grund muss ein Dialogsystem verschiedene Verfahren implementieren, um die sich in der Interaktionsgeschwindigkeit an den Benutzer anzupassen. Für SpeechActs entwickelten Yankelovich et al. diverse Methoden, den Gesprächsfluss zu beschleunigen.

Einmischen (Barge-In) - In den Testreihen versuchten Benutzer häufig erfolglos, bekannte Ansagen zu unterbrechen, um einen weiteren Befehl zu geben. Die Entwickler fügten daraufhin eine als "Barge-In" bezeichnete Funktionalität hinzu, die es dem Benutzer erlaubt, neue Befehle zu geben, während das System noch spricht. Die Sprachausgabe wird dann abgebrochen.

Freie Wahl der Wiedergabegeschwindigkeit - Eine weitere Methode um den Dialogfluss zu beschleunigen ist, die Geschwindigkeit der Sprachausgabe zu erhöhen. Damit können Benutzer bekannte Ansagen oder unwichtige Nachrichten schneller und wichtige Informationen langsamer abspielen.

Komprimierte Darstellung Um die Geschwindigkeit von Interaktionen zu erhöhen, sollten Ansagetexte möglichst kurz ausfallen. Dies kann unter anderem geschehen, indem eine Antwort mit dem nächsten Gesprächsschritt verbunden wird. Auch eine ausklingende Darstellung (tapered presentation) ist sinnvoll. Dabei werden sich wiederholende Satzteile weggelassen:

Momentan wird Sun Microsystems bei 25 gehandelt, einen halben Punkt mehr als gestern. Silicon Graphics liegt bei 23 1/2, um ein Halb gefallen. IBM bei 69, ein Achtel gestiegen.

Sprachkonventionen Beschleunigend auf den Gesprächsverlauf wirkt sich auch die Nutzung von Sprachkonventionen aus, um dem Benutzer mitzuteilen, was er als nächstes sagen könnte:

SpeechActs: "*Sie haben heute drei Termine*".

Benutzer: "Und was habe ich morgen?"

Tastaturkürzel (Shortcuts) - Besonders im Vergleich zu Systemen, die Eingaben über die Wähltastatur entgegennehmen wirken Sprachdialogsysteme oft langsam. Tastaturkürzel für häufig benutzte Aktionen beschleunigen hier die Arbeit und reduzieren gleichzeitig die Belastung für die Stimme.

Audio-Icons - Für fortgeschrittene Benutzer können zusätzlich "Audio-Icons" implementiert werden, die lange, gesprochene Ansagen ersetzen. Anstelle der Ansage "Lösche E-Mail" kann dann ein Soundclip mit einem "Papier-zerkrüllen"-Geräusch treten.

Geschwindigkeit des Systems optimieren - Bei der Implementierung von SpeechActs 1995 war die Hardware noch nicht leistungsfähig genug, um die Spracherkennung in Echtzeit durchzuführen. Dies resultierte in einem etwas langsameren Gesprächstempo, das den Benutzern als störend und unnatürlich auffiel. Dieses Problem ist mit heutiger Hardware nicht mehr gegeben.

Vokabular Bei grafischen Anwendungen muss der Benutzer sich - mal mehr, mal weniger - an die Eingabe- und Ausgabemöglichkeiten der grafischen Benutzerschnittstelle anpassen. Im Gegensatz dazu erwartet der Benutzer, dass ein Sprachdialogsystem sein Vokabular versteht und spricht.

Idealerweise lehnt sich das Vokabular des Sprachdialogsystems an das zwischenmenschlicher Gespräche an. Dabei reicht es aber nicht aus, alle Äußerungen des Benutzers zu verstehen. Vielmehr muss der Computer implizite, kontextbezogene Informationen auswerten können.

Ein grafischer Kalender hat kein Konzept für den Begriff "nächsten Montag". Der Benutzer übersetzt dies selbst in ein entsprechendes Datum, das er dem Kalender mitteilt. Vom Dialogsystem hingegen wird erwartet, dass es diese Information selbständig umwandelt. Aber nicht nur Daten, auch Befehle sind kontextsensitiv und können sich auf vorhergehende Antworten beziehen:

Benutzer: "Welche Termine habe ich morgen?"

SpeechActs: "... "

Benutzer: "Und am Freitag?"

Trotz der oben beschriebenen Bemühungen, die Interaktion möglichst stromlinienförmig und kompakt zu gestalten, verwendet SpeechActs einige Standardfloskeln wie "Auf Wiedersehen" am Ende eines Dialogs. Dies weicht die knappe Sprache etwas auf und nähert die Dialoge an reale Konversation an.

Umwandlung von grafischen in Sprach-Benutzerschnittstellen Im Gegensatz zur Interaktion mit grafischen Benutzerschnittstellen fehlt es dem Benutzer von Sprachdialogsystemen an visuellem Feedback. Bei einer Grafikanwendung weiß er jederzeit, in welchem Kontext oder Modus er sich gerade befindet. Da Sprache allerdings eine sehr stark temporale Natur hat, vergisst der durchschnittliche Benutzer schnell, an welcher Position im Dialog er sich gerade befindet.

Neben der Herausforderung, die menschliche Konversation zu imitieren, ist es daher besonders wichtig, dass der Benutzer sich ein mentales Modell der Interaktionsmöglichkeiten bilden kann. Dazu gehört, dass er weiß, wo im Dialog er sich gerade befindet und was seine Optionen im Moment sind. Diese Designherausforderung manifestiert sich in vielen Details:

Kontinuität Besonders verwirrend ist, wenn sich der Dialogkontext vom Benutzer unbeabsichtigt ändert.

Im SpeechActs-System trat dieses Problem beispielsweise im Posteingang auf: Im grafischen E-Mail-Programm das bei Sun verwendet wird sind die Nachrichten im Posteingang durchnummeriert. Wird dort eine E-Mail als gelesen markiert, verschwindet sie aus dem Posteingang und die restlichen Nachrichten werden wieder von vorne durchnummeriert, so dass keine Lücken entstehen.

Auch im SpeechActs-System konnte der Benutzer auf Nachrichten im Posteingang über deren laufende Nummer zugreifen. Wenn eine E-Mail als gelesen markiert wurde und aus dem Posteingang verschwand, änderten sich auch hier

für die anderen Nachrichten die zugeordneten Nummern. Da der Benutzer aber bei Sprachdialogsystemen die Betreffzeilen der Nachrichten schlechter scannen kann als in einem grafischen E-Mail-Client, ist er stark darauf angewiesen, auf eine Nachricht über ihre Nummer zugreifen zu können.

Yankelovich et al. lösten dieses Problem, indem sie gelesene Nachrichten nicht sofort aus dem Posteingang entfernen. Stattdessen bekommen diese ein "Gelesen"-Attribut gesetzt. Wenn der Benutzer später wieder auf die bereits gelesene Nachricht zugreift, informiert ihn SpeechActs lediglich darüber, dass er diese Nachricht bereits gelesen hat. Erst bei Beendigung einer Sitzung werden die gelesenen Nachrichten aus dem Posteingang entfernt.

Subdialoge Grafische Benutzerschnittstellen verwenden oft Dialogfenster und Assistenten (Wizards), um komplexere, eigenständige Aufgaben auszuführen. Auch in SpeechActs kommen solche Subdialoge vor, zum Beispiel, wenn der Benutzer auf eine E-Mail antworten will. Diese sind allerdings nicht so deutlich von der restlichen Benutzerschnittstelle abgehoben, wie in grafischen Benutzerschnittstellen. Ziel war es nun, dem Benutzer mitzuteilen, wenn er einen Subdialog erfolgreich abgeschlossen hatte und wieder zur normalen Programmoberfläche zurückkehrte. Die Sun-Forscher versuchten es erst mit einem Signalton, wenn die Antwort verschickt wurde. Dies wurde dann zugunsten eines gesprochenen "Was nun?" fallengelassen.

Status-Signalisierung Wenn das System nach einem Befehl schweigt, weiß der Benutzer nicht, ob es gerade arbeitet, den Befehl nicht verstanden hat oder sogar abgestürzt ist. Um die Tätigkeit des Systems hörbar zu machen, kann - ähnlich der Sanduhr in grafischen Benutzeroberflächen - ein Warteklang oder eine Wartemelodie abgespielt werden.

Dialogsteuerung durch den Computer Während grafische Benutzeroberflächen dem Benutzer oft einen festen Weg vorgeben, um eine Aufgabe zu erfüllen, muss sich eine Sprach-Benutzerschnittstelle dem Benutzer unterordnen, um akzeptiert zu werden.

Ein häufig vorkommendes Verhalten bei Benutzern von Sprachdialogsystemen ist, nicht immer explizit auf Fragen zu antworten. Auf die Frage "Ihre Nachricht wird an Matt Marx verschickt. Okay?" antworteten im Test viele entweder gleich mit einem neuen Befehl ("Nächste Nachricht") oder hängten den nächsten Befehl an die Antwort an ("Ja, nächste Nachricht"). Dieses Verhalten wurde auch von anderen Forschern festgestellt [5].

SpeechActs behilft sich hier mit einer toleranten Strategie: Als Antwort auf eine Frage sind auch "Ja, ..." und "Nein, ..." erlaubt. Wenn keine explizite Bestätigung mit "Ja" oder "Nein" erfolgt, macht SpeechActs "das Richtige", das heißt, das Dialogsystem führt dann automatisch die Aktion aus, die am wenigsten destruktiv ist. Wenn wichtige Fragen explizit beantwortet werden sollen, greift SpeechActs auf anleitende Aufforderungen (Directive Prompts) zurück, die dem Benutzer die Antwort vorschreiben: "Sagen Sie Ja, Nein oder Abbrechen!"

Erkennungsfehler Das größte Problem von Sprachdialogsystemen sind die Mängel der Spracherkennungssoftware. Dadurch kommt es zu inkonsistentem Verhalten, weil der Computer einmal einen Befehl versteht, beim nächsten Mal aber nicht.

Denn wenn der Benutzer einen Befehl gibt und eine Aktion ausgeführt wird, stellt er einen Ursache-Wirkung-Zusammenhang her. Ein inkonsistentes Verhalten des Rechners verhindert somit, dass der Benutzer sich ein konzeptuelles Modell (conceptual model) baut, wie die Anwendung reagieren wird.

Generell werden drei Arten von Erkennungsfehlern unterschieden:

Zurückweisung (Rejection Error) Das Dialogsystem kann keine gültige Eingabe erkennen und verweigert den Dienst. Dieser Mauereffekt (brick wall effect) führt sehr schnell zu Frustration beim Benutzer. Abhilfe schafft hier eine schrittweise Hilfestellung (progressive assistance): Beim ersten Erkennungsfehler gibt das System eine kurze Fehlermeldung aus, wenn der Fehler wiederholt gemacht wird bietet das System ausführlichere Hilfestellung an.

SpeechActs: Was haben Sie gesagt?

SpeechActs: Wie bitte?

SpeechActs: Wie bitte? Formulieren Sie die Aussage bitte um..

SpeechActs: Ich verstehe Sie nicht. Sprechen Sie deutlich, ohne übermäßig zu betonen.

SpeechActs: Es klappt immer noch nicht. Sprechen Sie bitte nach dem Piepton.

Wenn Fehler wiederholt auftreten, ist das oft ein Zeichen für Äußerungen, die nicht im Vokabular des Systems enthalten sind. Deshalb werden die Benutzer aufgefordert, ihren Befehl umzuformulieren. Die progressive Hilfestellung sorgt damit auch dafür, dass der Benutzer die richtigen Befehle lernt.

Ersetzung (Substitution Error) Die Zurückweisung von Befehlen ist "nur" störend, wohingegen das falsche Verständnis eines Befehls großen Schaden anrichten kann. Zum Beispiel fragt der Benutzer vielleicht nach dem Wetter am Niederrhein, das System versteht aber "Auf Wiedersehen" und beendet die Sitzung. Deshalb ist es sinnvoll, den Benutzer bei kritischen Aktionen noch einmal um Bestätigung zu bitten. Um den Benutzer nicht übermäßig mit Rückfragen zu belästigen, kann man wieder zwei Fälle unterscheiden: Wird ein Befehl als Frage nach Auskunft verstanden ("Wie ist das Wetter in Athen?"), kann die richtige Erkennung implizit in der Antwort verifiziert werden. Dazu baut das Dialogsystem Teile der Frage in die Antwort ein ("Das Wetter in Athen ist sonnig bei 24 Grad."). Wird hingegen ein kritischer Befehl erkannt, der den Verlust von Daten zur Folge haben kann, fragt das System noch einmal explizit rück beziehungsweise entscheidet sich im Zweifelsfall für die nicht-destruktive Antwort.

Einfügen (Insertion Error) Hintergrundlärm oder vom Benutzer erzeugte Geräusche können von der Spracherkennung als Befehle aufgefasst werden. Dieser

Gefahr kann mit den oben beschriebenen Verfahren begegnet werden. Weiterhin kann ein "Stummschalten"-Knopf eingeführt werden, der die Spracherkennung abschaltet.

3.4 Schlüsse

Der SpeechActs-Prototyp hat ein grundlegendes Prinzip für das Design von Sprach-Benutzerschnittstellen etabliert: Es ist immens wichtig, die Kommunikation zwischen Mensch und Computer an die Grundlagen der zwischenmenschlichen Konversation anzugleichen. Wichtig ist auch, dass Benutzerschnittstellen für Sprache nicht einfach aus grafischen Benutzerschnittstellen abgeleitet werden können, sondern fast gänzlich eigenen Regeln unterliegen.

Bei der Spracheingabe tritt als limitierendes Element die unausgereifte Spracherkennung auf. Bei einer Erkennungsrate von 90% wird immer noch jeder zehnte Befehl falsch erkannt. Insofern sind Rückmeldungen und Verifizierung von Eingaben sehr wichtig. Um den Benutzer nicht zusehr zu belästigen, sollte die Verifizierung allerdings meistens implizit erfolgen ("Wie ist das Wetter in Athen?" - "Das Wetter in Athen ist sonnig bei 24 Grad."). Nur bei kritischen Aktionen wie dem Löschen einer Nachricht sollte das System explizit nachfragen.

4 Nomadic Radio: Sprache als Benutzerschnittstelle für mobile Endgeräte

Nitin Sawhney und Chris Schmandt stellen mit dem Nomadic Radio ein mobiles Endgerät vor, das den Benutzer auf Schritt und Tritt begleitet und ihn sowohl interaktiv als auch selbständig mit gesprochenen Informationen versorgt. Diese umfassen E-Mails, Nachrichten auf dem Anrufbeantworter oder Radionachrichten. Das Ziel des Nomadic Radio ist es, durch intelligente Benachrichtigungsstrategien den Benutzer möglichst wenig bei seiner eigentlichen Tätigkeit zu stören. Dafür setzt es neben der Sprache weitere auditive Elemente ein.

4.1 Grundlagen und Probleme der mobilen Kommunikation

Grundlegendes Problem beim Design eines mobilen Endgerätes ist die notwendige Vielseitigkeit der Benutzerschnittstelle. Benutzer von mobilen Kommunikationslösungen wie dem Nomadic Radio wollen unterwegs sowohl schnelle, kurze als auch ausgedehnte Transaktionen durchführen.

Sawhney und Schmandt unterscheiden außerdem zwei primäre Grundfunktionalitäten eines mobilen Endgerätes, Navigation und Notifikation. Zum einen will der Benutzer aktiv mit dem Gerät interagieren, zum anderen soll das System auch selbständig Benachrichtigungen ausgeben können.

Eine Lösung für diese Anwendungsszenarien sind auditive Schnittstellen.

Vorteile von Audio in mobilen Endgeräten Im Vergleich zu grafischen Benutzerschnittstellen haben auditive Benutzerschnittstellen gewisse Vorteile in mobilen Endgeräten.

- Mobile Endgeräte werden zunehmend kleiner. Bei diesen Größen stoßen grafische Benutzerschnittstellen an die Grenzen der Bedienbarkeit. Bildschirm und Tastatur verlieren ihren Nutzen, da sie nicht beliebig verkleinert werden können.
- Sprachein- und -ausgabe ermöglichen eine unauffällige Bedienung des Systems. Hände und Augen sind frei für andere Dinge.
- Durch einfache, zuverlässige Sprachkommandos können Aktionen schneller ausgeführt werden, als mit langwierigem Tippen von Befehlen.
- Gesprochene Sprache ist ausdrucksstärker und effizienter als Schrift und beansprucht die Aufmerksamkeit des Benutzers weniger.
- Der Mensch kann gleichzeitig mehrere Audioquellen parallel und unterbewusst wahrnehmen und seine Wahrnehmung gezielt auf einzelne richten.

Problemstellungen

- Sprache kann ermüden und langsam sein. Deshalb sollte der Benutzer Sprachausgaben beschleunigen und überspringen können.
- Im Gegensatz zu grafischen Benutzerschnittstelle muss der Benutzer die möglichen Befehle vorher auswendig lernen.
- Ungünstige Umgebungseinflüsse können die Spracherkennung stören. Deshalb empfiehlt es sich, zusätzlich auch die Eingabe über Tasten zu ermöglichen.
- In unserer Gesellschaft gibt es soziale Konventionen. Es wirkt unnatürlich, wenn jemand scheinbar mit sich selbst spricht.
- Sprache ist öffentlich. Privatsphäre und Vertraulichkeit von Daten sind schwer zu wahren.

4.2 Das Nomadic Radio

Das Nomadic Radio ist ein Prototyp eines mobilen Kommunikations- und Informationssystems. Als Ein- und Ausgabe-Medium kommt ein VoiceBeam Neckset (Abbildung 2) zum Einsatz. Es ist bequem zu tragen und verfügt über ein Mikrofon und Stereo-Lautsprecher.

Die Java-basierte mobile Anwendung läuft auf tragbaren Rechnern (wearable computers). In den Programmiersprachen C und Perl geschriebene Server zeichnen Nachrichten aus dem Radio auf und bereiten E-Mails und Sprachnachrichten auf. Server und Client kommunizieren über WLAN. Die Sprachanalyse und Sprachsynthese übernimmt die AT&T Watson Speech API. Die Erkennungsrate beträgt über 90%, sinkt aber in lauten Umgebungen auf 50%. Der Benutzer kann während er einen Push-To-Talk(PTT)-Knopf drückt dem Nomadic Radio Befehle geben. Alternativ kann das System auch kontinuierlich lauschen. Eine Befehlssitzung wird dann mit "Nomadic Wake Up" eingeleitet. Wenn der Benutzer seine Sitzung beenden will, versetzt er das System mit "Nomadic Sleep" wieder in den Lauschzustand.



Abbildung 2. Voicebeam Neckset als Ein- und Ausgabemedium für das Nomadic Radio, [4]

4.3 Verwendete Techniken

Das Nomadic Radio implementiert eine Reihe raffinierter Techniken um die Aufmerksamkeit des Benutzers nicht zu sehr zu belasten.

Räumliches und paralleles Hören Die räumliche Anordnung von Geräuschquellen vermittelt dem Benutzer Informationen über die Art der Nachricht und erlaubt es ihm, selektiv einzelne Audiostreams herauszuhören. Das VoiceBeam Neckset hat Stereo-Lautsprecher, über die ein simulierter 3D-Surround-Sound ausgegeben wird. Empfangene Nachrichten werden nach Ankunftszeit um den Kopf herum angeordnet. Eine Mittags erhaltene Nachricht erklingt direkt von vorne, während eine Nachricht, die um 16 Uhr ankommt von rechts hinten zu hören ist. Der Benutzer kann sich auf eine Quelle konzentrieren und die anderen Nachrichten unterbewusst im Hintergrund wahrnehmen. Ein weiteres Feature, das die Fähigkeit des Menschen zur räumlichen Audiowahrnehmung ausnutzt, ist die Nachrichtenvorschau (Message Scanning). Das System liest den Anfang einer neuen Nachricht vor, wartet kurz, ob der Benutzer auf diese zugreifen will und spielt ansonsten den Anfang der nächsten Nachricht ab. Die Audioquellen wandern dabei lauter werdend von rechts zur Mitte. Anschließend klingen sie nach links aus, während die nächste Nachricht schon rechts erklingt. Weil der Mensch sich auf eine räumlich angeordnete Quelle konzentrieren kann, kann der zeitliche Abstand zwischen zwei Nachrichten reduziert werden, was die Vorschau erheblich beschleunigt.

Skalierbare Audiodarstellung Heutige mobile Endgeräte wie Mobiltelefone oder PDAs benachrichtigen den Benutzer auch zu unpassenden Gelegenheiten über Anrufe und Termine. Dies kann während einer Besprechung passieren, aber auch während der Benutzer sich stark konzentrieren muss. In den meisten Fällen rechtfertigt es die Priorität einer Nachricht nicht, den Benutzer dabei zu stören. Unterbrechungen kosten Zeit, laut einer Studie von O'Conaill und Frohlich 1995 [7] bei Büroarbeitern zehn Minuten pro Stunde.

Probleme von heutigen Systemen

- Undifferenzierte Signaltöne vermitteln nur binäre Informationen (Anruf / kein Anruf)
- Mangels Kenntnis der Umgebung des Benutzers wird dieser oft zu unpassenden Zeiten unterbrochen.
- Die Geräte lernen nicht aus dem Verhalten des Benutzers und belästigen ihn immer wieder mit den gleichen Problemen.
- Das System hat keinen Überblick über die empfangenen Nachrichten. Dies führt zu unkoordinierten Benachrichtigungen.

Ein mobiles System sollte eine Benachrichtigungsstrategie verfolgen, die den Kontext des Benutzers miteinbezieht und zum Beispiel unwichtige Nachrichten erst später abspielen, wenn der Benutzer gerade Zeit hat. Deshalb muss es ein Bild der Umgebung haben (zum Beispiel durch passives Mithören der Umgebungsgeräusche) und auch die Wichtigkeit einer Nachricht und die Aktivität des Benutzers miteinbeziehen.

Das Nomadic Radio setzt die Forderung nach einer kontextabhängigen Benachrichtigung mit einer Skalierbaren Audiodarstellung (Scaleable Auditory Presentation) um.

Je nach Situation und Priorität der Nachricht stellt das System diese durch unterschiedliche Signale dar. Ziel ist, die Informationen dem Benutzer möglichst unobtrusiv darzustellen.

Stille. Wenn der Benutzer absolut nicht gestört werden darf, bleibt das System stumm. Nur Nachrichten mit höchster Priorität werden ausgegeben.

Hintergrundgeräusche (Ambient Audio) wie der Klang von fließendem Wasser dienen als unaufdringliche Signalisierung von Netzaktivität. Wenn Daten übertragen werden fließt das Wasser schneller, ankommende Nachrichten werden als leise Platscher wiedergegeben.

Benachrichtigungs-Klänge teilen die Dringlichkeit einer Nachricht mit. Der Benutzer kann bestimmten Gruppen eigene Klänge zuordnen.

VoiceCues sind 1-2 Sekunden kurze charakteristische Stimmproben. Wenn eine neue E-Mail eintrifft, ertönt die Stimme des Absenders. Dies hat sich als besonders effektive Informationsquelle erwiesen, da die menschliche Stimme viel Aufschluss über die Dringlichkeit einer Nachricht geben kann, leicht einem Absender zuzuordnen ist und .

4.4 Evaluation und Schlüsse

Zwei Testpersonen verwendeten ein Nomadic Radio einige Tage für ihre üblichen Messaging-Aktivitäten. Sie hatten Probleme, das System zu benutzen, wenn Sie mit anderen Menschen zusammen waren, und bevorzugten auch die "Hintergrundgeräusche" gegenüber gesprochenen Benachrichtigungen. Des weiteren wünschten sich beide eine Zweitastenbedienung (OK, Abbrechen) zusätzlich zur Sprach-Benutzerschnittstelle, um reine Bestätigungen schneller ausführen zu können.

Auf mobilen Endgeräte ermöglichen auditive Benutzerschnittstellen eine kompakte und benutzerfreundliche Kommunikationsmethode. Wichtig ist allerdings, den Benutzer nicht mit Informationen zu überfluten. Unterschwellige Benachrichtigungen durch Hintergrundgeräusche bieten sich deshalb an und werden auch gerne angenommen.

5 Grundregeln für Sprach-Benutzerschnittstellen

Auch für das Design von Sprach-Benutzerschnittstellen gibt es allgemeine Empfehlungen. Steven Ross, Elizabeth Brownholtz und Robert Armes fassen ihre Erfahrungen im Design von Benutzerschnittstellen in *Voice User Interface Principles for a Conversational Agent* zusammen.

5.1 Anpassung an menschliche Sprachgewohnheiten

- Das System muss natürliche Sprache verstehen und bei Erkennungsproblemen nachfragen
- Das System muss Höflichkeitsformen wie "Guten Morgen" oder "Danke" verstehen.
- Das System darf Benutzer nur bei wichtigen Benachrichtigungen unterbrechen. Der Benutzer darf das System immer unterbrechen.
- Das System soll nachfragen, ob es sprechen darf.
- Der Benutzer darf nicht in einem Modus feststecken und soll keine gestellten Fragen beantworten müssen, um einen anderen Befehl geben zu können.

5.2 Verlässlichkeit und vertrauensbildende Maßnahmen

- Bei Antworten soll das System die ursprüngliche Frage mit einbauen ("Das Wetter in Athen ist sonnig bei 24 Grad Celsius" statt "Sonnig bei 24 Grad Celsius")
- Wenn der Erfolg einer Aktion nicht implizit wahrnehmbar ist, muss das System eine Bestätigung ausgeben.
- Destruktive oder nicht reversible Aktionen muss der Benutzer explizit bestätigen
- Bei länger andauernden Operationen muss das System deutlich machen, ob und was es macht.
- Der Benutzer sollte Operationen unterbrechen können.
- Wenn der Benutzer keine Antwort auf eine Frage gibt, sollte das System die Frage wiederholen.

5.3 Konsistenz und Transparenz des mentalen Modells

- Das System sollte nach Möglichkeit die Sprache verwenden, die es vom Benutzer erwartet. Es sollte keine Worte benutzen, die es nicht selbst versteht.

- Das System sollte keine Vermutungen über die Absichten des Benutzers anstellen, sondern nur das tun, was ihm explizit befohlen wird.
- Die Sprachbenutzerschnittstelle sollte konsistent sein, gleiche Aktionen sollten auf die gleiche Weise ausgeführt werden. Ähnliche Formulierungen sollten zusätzlich zur "konsistenten" Form erlaubt sein.
- Das System sollte den Benutzer nicht glauben machen, es sei intelligenter als es ist, da sonst der Benutzer Befehle gibt, die das System überfordern.

6 Sprache als Geräusch: Nonverbale Eingabemöglichkeiten

Neben Sprache kann man auch andere Fähigkeiten der menschlichen Stimme nutzen, um mit dem Computer zu kommunizieren. Takeo Igarashi und John F. Hughes von der Brown University haben drei Interaktionstechniken untersucht, die sie als "Control by Continuous Voice", "Rate-based Parameter Control by Pitch" und "Discrete Control by Tonguing" bezeichnen.

"*Control by Continuous Voice*" ermöglicht es, Parameter durch die Länge eines anhaltenden Tons zu beeinflussen. Der Benutzer sagt dann zum Beispiel seinem Fernseher "Lauter, aaaaaaaaaaaaaah". Solange der Benutzer einen bestimmten Ton von sich gibt, wird die Lautstärke kontinuierlich erhöht. Ein ähnliches System existiert mit SUITEkeys [8]. Dieses erfordert aber keinen kontinuierlichen Ton sondern einen Stopp-Befehl ("Maus nach unten stopp").

"*Rate-based Parameter Control by Pitch*" erweitert das oben beschriebene Verfahren um die Möglichkeit, einen zweiten Parameter durch die Tonhöhe zu steuern. Im obigen Szenario könnte der Benutzer die Tonhöhe variieren, um die Geschwindigkeit der Parameteränderung zu beeinflussen. Je höher der Ton, desto schneller wird die Lautstärke erhöht.

"*Discrete Control by Tonguing*" ermöglicht die Eingabe von diskreten Werten. Dazu macht der Benutzer für jeden Schritt ein Geräusch: Mit "Kanal hoch, ta, ta, ta, ta" kann er den Fernsehsender wechseln.

Diese Verfahren bieten eine sprachunabhängige, einfache und sehr direkte Steuerung von einfachen Parametern. Im Unterschied zu Sprachbefehlen ermöglichen Sie nuancierte und gleichzeitig kurze Befehle. Für den menschlichen Stimmapparat ist es aber relativ anstrengend, längere Zeit solche Geräusche hervorzubringen. Dies schränkt die Anwendung auf kurze, nicht zu häufige Befehlssequenzen ein.

7 Diskussion

Obwohl alle vier Arbeiten unterschiedliche Schwerpunkte haben, kristallisieren sich folgende drei grundlegenden Eigenschaften einer guten Sprach-Benutzerschnittstelle heraus:

Stringenz Sprechen ist naturgemäß langsamer und ermüdender als das Tippen auf einer Tastatur. Aufgrund der temporalen Natur von Sprache muss der Benutzer Sprachausgaben des Systems in voller Länge abhören und kann sie nicht wie bei grafischen Benutzerschnittstellen einfach ignorieren oder überfliegen. Deshalb ist es wichtig, den Benutzer zum einen von wiederkehrenden Sprachbefehlen zu entlasten und diese durch Befehltasten zu ersetzen. Außerdem sollte ein Sprachdialogsystem die Möglichkeit bieten, die Audioausgabe zu beschleunigen, zu verlangsamen und zu überspringen.

Konsistenz Im Gegensatz zur Benutzung einer grafischen Benutzerschnittstelle weiß der Anwender bei einer Sprach-Benutzerschnittstelle nicht jederzeit, in welchem Modus er sich befindet und welche Befehle ihm zur Verfügung stehen. In den meisten Anwendungsfällen hat der Benutzer nicht die Möglichkeit, ein Handbuch zu konsultieren. Aus diesem Grund sollte das Vokabular der Applikation einfach sein und konsistent bleiben. Das System muss dem Benutzer bei Bedarf Hilfestellung bieten. Es kann sinnvoll sein, je nach Aktion oder Modus wechselnde Hintergrundgeräusche zu verwenden, die dem Benutzer eine Orientierungshilfe geben.

Flexibilität Benutzer verwenden Synonyme für Befehle oder versuchen neue Befehle. Dies sollte das System berücksichtigen und Hilfestellung anbieten, wenn es eine Eingabe nicht verstehen kann. Auch Höflichkeitsformen, Umschreibungen und Äußerungen wie "Äh" und "Moment..." dürfen das System nicht durcheinanderbringen.

Besonders bei der Verwendung in in mobilen Geräten ist eine kontextsensitive Benachrichtigungsstrategie notwendig, die Störungen minimiert. Räumliche Audiodarstellung bietet die Möglichkeit, neben der eigentlichen Nachricht auch noch Metadaten übertragen zu können. Das Nomadic Radio und die Laut-Benutzerschnittstelle von Takeo Igarashi und John Hughes zeigen, dass non-verbale Ausdrucksmöglichkeiten wie unterschwellige Hintergrundgeräusche oder Stimmlaute Sprach-Benutzerschnittstellen deutlich bereichern können.

8 Stand der Technik: Praktische Anwendungen von Sprach-Benutzerschnittstellen

Sprachdialogsysteme sind heute nicht mehr wegzudenken. Vor allem im für telefonische Auskunftsdienste und Kundenbetreuung verwenden viele Unternehmen die künstlichen Kommunikationspartner. Da die Entwicklung von Software für Spracherkennung, -analyse, -generierung und -ausgabe aber hohen Forschungs- und Programmieraufwand erfordert, ist der Markt für Sprachtechnologie überschaubar. Im Folgenden werden einige kommerzielle und Open-Source-Applikationen für Sprachanwendungen vorgestellt. Ein weiterer Abschnitt beschäftigt sich mit VoiceXML, einer standardisierten Beschreibungssprache für Sprach-Benutzerschnittstellen.

8.1 Enterprise-Software

Viele Anbieter für Sprachapplikationen im Unternehmensbereich bieten ein umfassendes Portfolio aus Software für Spracherkennung und Sprachausgabe sowie kompletten Sprachdialogsystemen. Marktführer sind die beiden US-amerikanischen Unternehmen Scansoft und Nuance. In Deutschland existieren mit Crealog, Sympalog und Clarity ebenfalls einige Firmen mit eigenentwickelten Sprach-Portalen. Das US-Unternehmen Wizzard vertreibt die Text-To-Speech-Engine *AT&T Natural Voices*, die hochwertige Stimmen in vielen verschiedenen Sprachen bietet.

8.2 Anwender-Software

Für Büroanwender und Privatpersonen gibt es seit einiger Zeit schon Software zur Spracherkennung. Am weitesten verbreitet sind *Dragon Naturally Speaking* und IBM *ViaVoice*, dessen Engine auch in Linguatecs *Voice Pro* integriert ist. Beide Programme bieten spezielle Versionen für Mediziner und Juristen, die sich durch Wörterbücher mit Fachbegriffen auszeichnen. Der Webbrowser Opera verwendet IBM-Technologie um Sprachkommandos zu empfangen und Webseiten vorzulesen.

Schon seit längerem existieren rudimentäre Spracherkennung und Sprachausgabe auch in den Betriebssystemen Windows und MacOS. Besonders die Windows-Engines, die optional mit dem Microsoft Agent installiert werden können, sind aber nur eingeschränkt verwendbar. Sowohl Microsoft als auch Apple fügen ihren nächsten Betriebssystemen aber neue Sprach-Funktionalität hinzu. Apple hat mit *VoiceOver* eine Text-to-Speech-Engine und eine Spracherkennung in Mac OS 10.4 (Tiger) integriert. Microsoft plant, im Windows-XP-Nachfolger Longhorn eine Sprach-API zu integrieren.

8.3 Open-Source-Software

Es existieren einige Open-Source-Anwendungen für Sprach-Benutzerschnittstellen. Diese stammen vorwiegend aus dem akademischen Bereich und laufen meistens nur unter Linux. Einige werden schon längere Zeit nicht mehr weiterentwickelt. CMS Sphinx [12] ist eine Software zur Spracherkennung, die in verschiedenen Versionen als C- und Java-Implementationen frei zur Verfügung steht. Einer weitere freie Spracherkennungssoftware ist ASR [17]. Zur Analyse der erkannten Sprache dient der Text-Parser Phoenix [11]. Zur Sprachausgabe verwenden viele Open-Source-Programme Festival [13]. Die TTS-Software existiert auch in einer abgespeckten Version für Embedded-Computer. Eine Alternative zu Festival stellt das tschechische Epos [15] dar. Als Plattform für Sprachdialogsysteme hat sich der Galaxy Communicator [14] der MIT Spoken Language Group etabliert. Auch mit dem VoiceXML-Browser publicVoiceXML [16] kann man einfache Dialogsysteme entwerfen.

8.4 VoiceXML

Was ist VoiceXML? VoiceXML ist eine Beschreibungssprache für Sprach-Benutzerschnittstellen. Der offene, auf XML basierende Standard wird von vielen Systemen unterstützt und findet seinen Weg auch in Anwenderprogramme wie den Webbrowser Opera. VoiceXML ist nicht auf freie Konversation mit dem Computer ausgelegt, sondern modelliert hauptsächlich Formulare, die durch Sprachbefehle ausgefüllt werden können.

Geschichte Der Ursprung von VoiceXML liegt in der Phone Markup Language (PML), die 1995 im Rahmen eines AT&T-Projekts entstand, das den Designprozess von Spracherkennungssoftware vereinfachen sollte. Auch Motorola, Lucent und IBM arbeiteten an eigenen PML-Versionen. 1998 gründeten AT&T, IBM, Motorola und Lucent das VoiceXML Forum, um an der Standardisierung der Sprache zu arbeiten. 2000 veröffentlichten sie VoiceXML 1.0, das sie kurz darauf beim W3C als Basis für einen internationalen Standard vorschlugen. 2004 wurde VoiceXML 2.0 als W3C Recommendation veröffentlicht.

Syntax VoiceXML verwendet zwei Arten von Datendateien. Die eigentlichen VoiceXML-Dateien modellieren die Benutzerschnittstelle, während die grXML-Dateien die Grammatiken für Elemente festlegen. Eine Grammatik legt dabei fest, welche Eingaben in ein bestimmtes Feld zulässig sind.

```
<vxml version="2.0">
  <form>
    <field name="drink">
      <prompt>Would you like coffee, tea, or nothing?</prompt>
      <grammar src="drink.grxml" type="application/srgs+xml"/>
    </field>
  </block>
  <submit next="http://www.drink.example.com/drink2.asp"/>
</block>
</form>
</vxml>
```

VoiceXML-Datei, die nach dem gewünschten Getränk fragt und das Ergebnis an ein ASP-Skript schickt.

```
<grammar mode="voice" xml:lang="en-US" root="command">
  <rule id="command" scope="public">
    <one-of>
      <item>coffee</item>
      <item>tea</item>
      <item>milk</item>
      <item>nothing</item>
    </one-of>
  </rule>
</grammar>
```

grXML-Datei, die festlegt, welche Antworten auf eine Frage gegeben werden können.

Wird keine zulässige Eingabe getätigt, wiederholt die Applikation die Frage. Neben Grammatik-Dateien für Spracheingabe gibt es auch eine Version für DTMF-Tastentöne.

Elemente von VoiceXML VoiceXML-Applikationen können zwei Basiselemente verwenden: *Formulare*, um Daten einzugeben und *Menüs*, um Aktionen auszuführen. Der Entwickler kann in VoiceXML auch *Sessions* und *Subdialoge* implementieren. Wenn Benutzereingaben nicht erkannt werden können und auch in anderen Sonderfällen werden *Events* ausgelöst, die von speziellen Methoden behandelt werden können.

Verbreitung von VoiceXML VoiceXML hat sich klar als Industriestandard durchgesetzt. Vor allem die großen Anbieter von Sprachdialogsystemen setzen auf VoiceXML zur Modellierung von Dialogen. Aber auch Anwendersoftware wie der Opera Browser und einige Open-Source-Applikationen [16] unterstützen VoiceXML.

9 Ausblick

Sprach-Benutzerschnittstellen bieten viele neue Möglichkeiten der Mensch-Maschine-Interaktion. Sie sind relativ billig in bestehende mobile Endgeräte zu integrieren und nutzen eine seit Jahrtausenden erprobte Interaktionform. In unserer mobilen Multitaskinggesellschaft erlauben Sprachtechnologien die ständige Verfügbarkeit von Informationen. Gleichzeitig können intelligente Benachrichtigungsstrategien uns vor dem Informationsinfarkt retten. Ob es aber Sinn des Fortschritts ist, dass wir immer mehr Sachen gleichzeitig tun können, sei dahingestellt. Humanistischer ist da schon die Vorstellung, dass die Sprache den Computer menschlicher und persönlicher macht. Statt, dass der Mensch sich an den Computer anpasst passt sich nun der Computer an den Menschen an. Aber auch körperlich behinderten Menschen und Kindern bietet die Sprachein- und -ausgabe eine hervorragende Alternative zu Tastatur und Maus.

Sprach-Benutzerschnittstellen werden weder grafische Benutzerschnittstellen noch Kommandozeilen ersetzen. Erstere haben ihnen gegenüber den Vorteil der besseren Übersicht und der steileren Lernkurve. Letztere sind der Spracheingabe in der Geschwindigkeit überlegen und bieten eine überragende Funktionalität. Statt vorhandene Benutzerschnittstellen zu ersetzen, erweitern Sprach-Benutzerschnittstellen diese vielmehr um einen weiteren Kanal, der besonders für kurze und räumlich variable Sitzungen geeignet ist. Spracherkennung erlaubt dann zum Beispiel in der Heimautomation, das Licht über Sprachkommandos zu steuern. Daneben werden vor allem mobile Endgeräte wie Mobiltelefone, PDAs und Spielkonsolen in Zukunft vermehrt auf Sprach-Benutzerschnittstellen setzen. Mit VoiceXML liegt auch ein anerkannter Standard vor, der die Entwicklung vorantreiben dürfte.

10 Literatur

1. Ross, S., Brownholtz, E., Armes, R.: Voice User Interface Principles for a Conversational Agent. Technical report, IBM Research (2004)
2. Igarashi, T., Hughes, J.F.: Voice as Sound: Using Non-verbal Voice Input for Interactive Control. In: UIST. (2001) 155–156
3. Yankelovich, N., Levow, G.A., Marx, M.: Designing SpeechActs: Issues in Speech User Interfaces. In: CHI. (1995) 369–376
4. Sawhney, N., Schmandt, C.: Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments. *ACM Transactions on Computer-Human Interaction* **7** (2000) 353–383
5. Kitai, M., Imamura, A., Suzuki, Y.: Voice Activated Interaction System Based on HMM-based Speaker-Independent Word Spotting. In: Proceedings of the Voice I/O Systems Applications Conference. (1991)
6. McTear, M.F.: Spoken Dialogue Technology: Enabling the Conversational User Interface. *ACM Computing Surveys* **34** (2002) 90–169
7. O’Conaill, B., Frohlich, D.: Timespace in the Workplace: Dealing with Interruptions. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI’95), New York, ACM (1995) 262–263

11 Links

8. SUITEKeys for Windows: (<http://www.suitekeys.org>)
9. Übersicht Dynamic Time Warping: (<http://www.dcs.shef.ac.uk/stu/com326/>)
10. Wikipedia: Hidden Markov Models: (http://en.wikipedia.org/wiki/Hidden_Markov_Models)
11. Phoenix - Semantic Parser: (<http://cslr.colorado.edu/~whw/>)
12. CMU Sphinx - Speech Recognition Engine: (<http://cmusphinx.sourceforge.net>)
13. Festival - Speech Synthesis System: (<http://www.cstr.ed.ac.uk/projects/festival/>)
14. Galaxy Communicator - Dialogue System Toolkit: (<http://communicator.sourceforge.net>)
15. Epos - Speech Synthesis System: (<http://epos.ure.cas.cz>)
16. Public VoiceXML - VoiceXML Browser: (<http://www.publicvoicexml.org>)
17. ISIP ASR - Speech Recognition Engine: (<http://www.isip.msstate.edu/projects/speech/>)

3D-Benutzungsschnittstellen, 3D-Scanning und Rekonstruktion

Marc Breisinger

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
breising@cip.ifi.lmu.de

Zusammenfassung Die vorliegende Arbeit befasst sich mit der Erfassung und Generierung räumlicher Daten für die Computerverarbeitung und deren Manipulation. In erster Instanz geht es dabei um dreidimensionales Scanning, die allgemeinen Prinzipien (primär Triangulation) und verschiedene Methoden wie Laserscanning, Verfahren mit strukturiertem Licht und Stereoskopie. Deren mögliche Implementationen werden im Hinblick auf das zu digitalisierende Objekt bzw. die jeweilige Anwendung betrachtet und verglichen. Weiter beschäftigen wir uns mit Benutzungsschnittstellen für den Umgang mit dreidimensionaler Rauminformation. Dabei geht es sowohl um Konstruktionswerkzeuge, die sich nicht-interaktiver Techniken bedienen, als auch um Echtzeitschnittstellen, mit welchen der Benutzer direkt Strukturdaten generieren und manipulieren kann.

1 Einleitung

Das Thema Raumstruktur, deren Erfassung und Weiterverwendung ist so alt wie Geographie und Astronomie. Auf dem Gebiet der Datenverarbeitung wird daran geforscht seit es Computer gibt. Dennoch macht man noch heute kontinuierlich Fortschritte in diesem Bereich und es sind bei weitem noch nicht alle Möglichkeiten ausgeschöpft. Verschiedene Herangehensweisen geben uns Werkzeuge in die Hand, besonders präzise, besonders schnell und/oder besonders kostengünstig an Rauminformationen zu kommen, was uns immer neue Anwendungsmöglichkeiten, vor allem im Bereich Virtual Reality, computerunterstützte Realität (Augmented Reality) und Mensch-Computer-Interaktion, eröffnet.

Was Präzision betrifft ist nach wie vor die Erfassung der Koordinaten mittels Lasertriangulation die erste Wahl, allerdings auch die teuerste und langsamste. Häufig muss die erforderliche Hardware auch auf einen bestimmten Anforderungsbereich zugeschnitten werden, was zusammen mit dem Kostenfaktor die allgemeine Anwendbarkeit sehr erschwert. Aufgrund der Beschränkung auf statische Objekte bleiben uns auch viele Anwendungsmöglichkeiten, wie z.B. die einer Echtzeit-Mensch-Computerschnittstelle bei dieser Herangehensweise verwehrt. Stereoskopie und Methoden mit strukturiertem Licht können wesentlich kostengünstiger und weniger speziell realisiert werden, sind aber v.a. was erstere

angeht oft weniger präzise und/oder störungsanfälliger. Allerdings stellen viele Anwendungen, wie zum Beispiel die Erkennung von Mimik und Gestik, diesbezüglich keine großen Anforderungen, und die Realisierung ist wesentlich weniger aufwendig. Auch werden wir bei der Methodik mit strukturiertem Licht, welche auch von Seiten der Hardware wesentlich niedrigere Ansprüche stellt, sehen, das hier einige Möglichkeiten existieren, wodurch dennoch ein hohes Maß an Präzision erreicht werden kann. Der Vollständigkeit halber sei erwähnt, dass es noch viele andere Vorgehensweisen gibt, wie z.B. das Schichtbildverfahren, welches vor allem in der Medizin (z.B. Computertomographie) Anwendung findet. Diese sollen hier allerdings aufgrund des methodischen Unterschieds zu den restlichen Herangehensweisen und aus Platzgründen nicht weiter betrachtet werden. Der erste Abschnitt dieser Arbeit wird sich mit den Prinzipien der auf Triangulation basierenden Verfahren beschäftigen.

Im zweiten Abschnitt betrachten wir Möglichkeiten der Gestaltung von dreidimensionalen Inhalten und der Umgang mit solchen Daten, die sich durch die zu Anfang erklärten Verfahren ergeben oder diese ergänzen. Bisher wird in dieser Hinsicht meist ausschließlich ein 3D-Modellierer, klassisch mit Maus und Tastatur am Bildschirm, herangezogen. Wie aber schon angedeutet bieten sich wesentlich mehr und, je nach Anforderung, wesentlich adäquatere Möglichkeiten der Erzeugung und Manipulation von Daten im virtuellen Raum. Neben dem Ausschauen oder plastischem Gestalten von realen Objekten, die anschließend gescannt und, eventuell automatisch, mit Logik erweitert werden, gibt es auch verschiedene Ansätze, das Material mit dem modelliert wird mit der Fähigkeit auszustatten, ihre Raumstruktur selbstständig an den Computer zu übermitteln wie z.B. in [7] beschrieben. Auch liegt bei der Bearbeitung von dreidimensionalen Daten die Verwendung von Werkzeugen nahe, die ebenfalls alle Freiheitsgrade des Raumes nutzen. Den Gedanken kann man soweit fortführen, dass virtueller und realer Raum idealerweise eins sind und man somit direkt ohne bewusste Translation mit den Daten arbeiten kann. Auch hierzu werden zwei Beispiele aus der aktuellen Forschung betrachtet.

2 Prinzipien der räumlichen Rekonstruktion

Um Informationen über räumliche Verhältnisse zu erhalten, vollzieht das menschliche Gehirn einen Vergleich der beiden Bilder, die es von den Augen erhält. Je weiter versetzt ein betrachteter Punkt auf dem einen Bild von seiner Position auf dem anderen Bild erscheint, desto näher ist er am Betrachter. Bei beinahe allen (insbesondere bei den im folgenden besprochenen) Verfahren der Akquisition von räumlichen Daten wird ein Verfahren benutzt, das dem Vorgehen des Gehirns sehr ähnlich ist: Die Triangulation. Diese wird im Folgenden aus mathematischer Sicht kurz erläutert, um danach bei konkreten Methoden zur Anwendung zu kommen.

Bei der Triangulation geht man von einer Basislinie s (Abb. 1) bekannter Länge \overline{AB} aus und misst den Winkel α des Vektors v_1 zu s , wobei man den Abstand \overline{BC} berechnen möchte. Diesen erhält man mit Hilfe von γ , dem Winkel zwischen v_1 und v_2 durch

$$\overline{BC} = \frac{\sin \alpha * \overline{AB}}{\sin \gamma}.$$

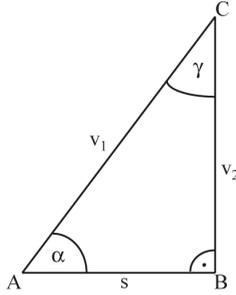


Abbildung 1. Triangulation.

Beim 3D-Scanning dient immer eine Kamera (O_1) als Sensor, eine zweite Kamera oder ein Projektor (O_2) als Referenz. In unserem Falle sind A und B die Brennpunkte der optischen Geräte O_1 und O_2 , deren Bilder miteinander verglichen werden. Die Vektoren v_1 und v_2 sind die Lichtstrahlen, die durch den jeweiligen Brennpunkt das reale Bildelement P mit dem zugehörigen virtuellen Element(Bildpunkt) verbinden. Da es in der Praxis schwierig ist γ zu messen, berechnen wir dieses mit $\gamma = \pi - (\alpha + \beta)$, wobei α und β aus den Abständen zu berechnen sind, den die korrespondierenden Pixel des aktuellen Bildpunkts von der Mitte des jeweiligen Bildes haben. Sei N der von uns gewählte Nullpunkt. Der Einfachheit halber setzen wir die Bildweite gleich dem Abstand e zwischen O_2 und N . Nun ist

$$\nu = \left(\arcsin \frac{\overline{AB}}{e} \right) * 2 \tag{1}$$

$$\delta_1 = \arctan(d_1/e) \tag{2}$$

$$\delta_2 = \arctan(d_2/e) \tag{3}$$

$$\alpha = (\pi - \nu)/2 + \delta_1 \tag{4}$$

$$\beta = (\pi - \nu)/2 - \delta_2 \tag{5}$$

$$\gamma = \pi - (\alpha + \beta) \tag{6}$$

$$\overline{BP} = \frac{\overline{AB} * \sin \alpha}{\sin \gamma} \tag{7}$$

$$z = e - \cos |\delta_2| * \overline{BP} \tag{8}$$

wobei z die gesuchte Tiefenkoordinate des Punktes P ist (siehe Abb. 2).

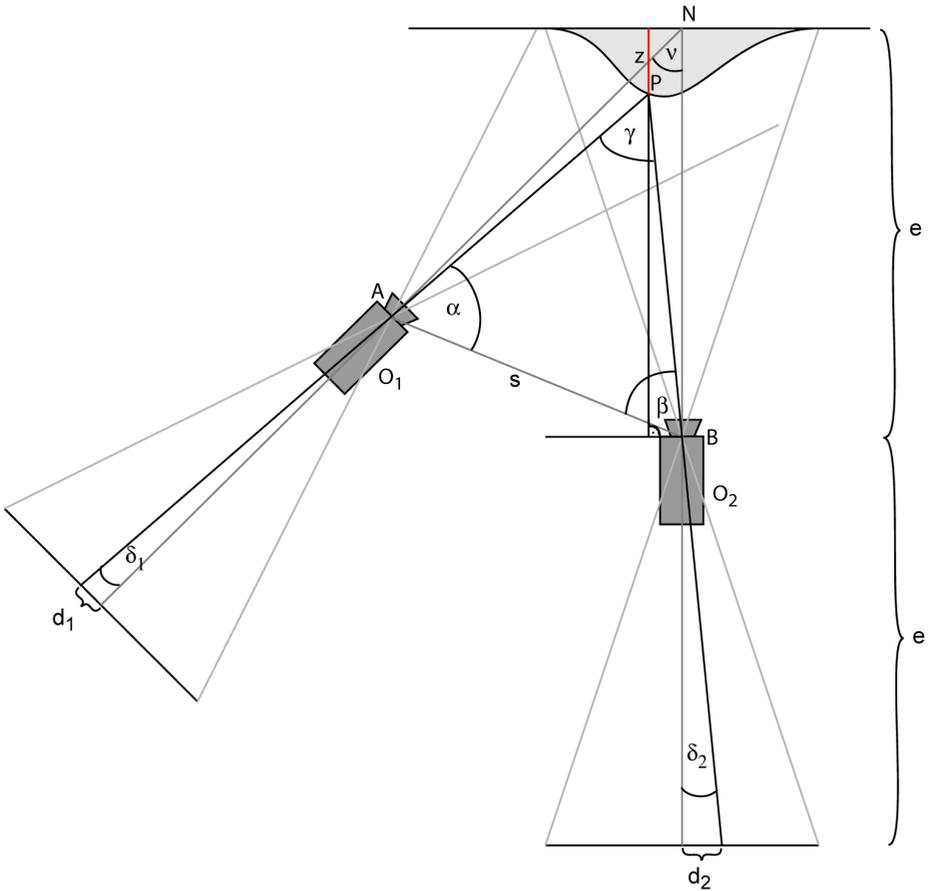


Abbildung 2. Berechnung der z-Koordinate des Bildelements P .

3 Methoden zur Datengewinnung

Die kritische Aufgabe beim Berechnen von Tiefeninformation ist nun, auf den zwei Bildern, die man nach dem Scan zur Verfügung hat, die korrespondierenden Bildpunkte zu identifizieren. Hierzu gibt es die verschiedensten Methoden die beliebig zuverlässig und komplex sein können. Die wichtigsten davon sollen hier beschrieben werden. Voraus angemerkt sei, dass es nicht unbedingt nötig ist entlang der Höhen- UND Breitenachse zu kodieren, da man einen Punkt erhalten kann indem man zwei Linien schneidet, oder auch indem man eine Linie (z.B. die Pixelreihe des Ausgangsbildes, die den gesuchten Punkt enthält) und eine

Ebene schneidet (z.B. die Ebene durch Brennpunkt von O_1 und Musterlinie. Siehe dazu 3.3). Allerdings kann ein Kodieren entlang beider Achsen zu Präzision und Robustheit beitragen.

3.1 Stereoskopie (Photogrammetrie)

In der Stereoskopie werden zwei Kameras benutzt, die das Objekt aus leicht unterschiedlichen Winkeln aufnehmen. In den erhaltenen Bildern wird dann nach korrespondierenden Bildpunkten gesucht. Da dieses Verfahren keinerlei Hilfsmittel verwendet, ist es störanfällig und algorithmisch relativ aufwändig. Kann man allerdings die Umgebungsbedingungen gut kontrollieren, wie z.B. in Abb. 4, hat dieses Verfahren einige Vorteile: Es dauert nur den Bruchteil einer Sekunde die erforderlichen Daten aufzunehmen und benötigt keine stark spezialisierte Hardware. Laut Siebert und Marshall[1] ist eine Präzision von bis zu 0.5 mm möglich. Sie benutzen zum Auffinden der korrespondierenden Bildpunkte einen Algorithmus, der aus den aufgenommenen Bildern eine Pyramide aus verschiedenen Auflösung erstellt. Die gröbste Auflösung bildet die Spitze, die feinste die Basis. Es wird nun beginnend bei der Spitze nach Übereinstimmungen gesucht (siehe Abb. 3) und sukzessive zur Basis durchgearbeitet, immer unter Zuhilfenahme der bisher gefundenen Erkenntnisse. Nachdem man auf Pixelebene angekommen ist, kann durch Interpolation Präzision auf Subpixelebene erreicht werden. Da sehr gleichmäßig texturierte Flächen, wie z.B. eine weiße Wand, dem Algorithmus Probleme bereiten, weil die Pixel kaum von einander zu unterscheiden sind, bietet es sich an, diese Methode mit einem Hilfsmittel aus dem Bereich Strukturiertes Licht (siehe 3.3) zu ergänzen. Hierfür wird während der Aufnahme ein Muster projiziert, dessen Eigenheiten leichter wieder zu erkennen sind.

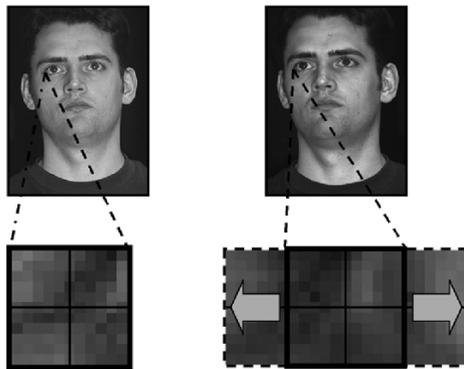


Abbildung 3. Auffinden von Korrespondenzen in Stereobildern [1].

Anwendungsbeispiel: Digitalisieren von Lebewesen Grundsätzlich erfordert das Scannen von beweglichen Objekten eine Methode, die mit einer sehr kurzen

Zeit für die Datenaufnahme auskommt. Außerdem hat die Photogrammetrie den Vorteil, keine eventuell augenschädigende Strahlung wie z.B. Laser zu verwenden. Sie bietet sich daher an um beispielsweise als Hilfsmittel in der Chiroplastik eine Simulation der geplanten Eingriffe zu ermöglichen, oder (in einer noch sehr stark in der Erforschung befindlichen Idee) für 3D-Filme. Letzteres scheidet momentan noch vor allem an dem enormen Datenaufkommen, dürfte sich aber schnell mit kommender Hardware weiterentwickeln lassen.

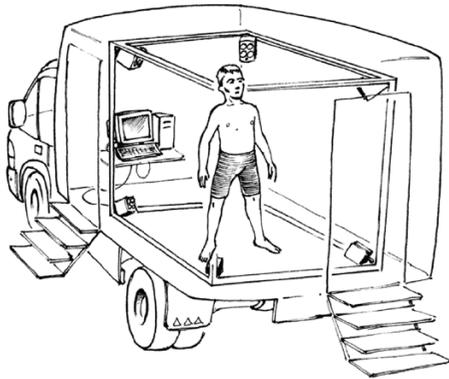


Abbildung 4. Setup für einen Ganzkörperscan [1].

3.2 Laserscanner

Laserscanner, die nach dem Triangulationsverfahren arbeiten (Alternativen sind Impulslaufzeitverfahren oder Phasenverschiebungsmessung) haben einen ähnlichen Aufbau. Anstelle der zweiten Kamera tritt hierbei allerdings ein Laseremitter (siehe Abb. 5). Der von dort ausgesandte Laserstrahl wird durch einen schnell oszillierenden Spiegel zu einem „Lichtfächer“ gestreut. Eine Kamera dient nach wie vor als Sensor. Sowohl ausgesandter Laserstrahl als auch die Lichtstrahlen, die zurück in die Kamera fallen, werden über einen Spiegel geleitet. Durch Schwenken dieses Spiegels streift der Laserstrahl, durch das Auffächern als Strich sichtbar, über das Arbeitsvolumen. Die Kamera macht nun Bilder von dem Laserstrich, der durch die Tiefenänderungen der Oberfläche verzerrt wird [14]. Das Ausmaß dieser Verzerrung ist die Grundlage für die oben besprochene Tiefenberechnung. Laut [10] und [8] ist eine Präzision im Hundertstelmillimeterbereich möglich.

Anwendungsbeispiel: Scannen statischer Objekte mit hoher Präzision Levoy et al. haben in ihrem Mammutprojekt, möglichst viele Statuen von Michelangelo mit Submillimeterpräzision zu digitalisieren, den Laserscanner gewählt, weil er die beste Kombination aus Präzision, Arbeitsvolumen, Störungsunempfindlichkeit

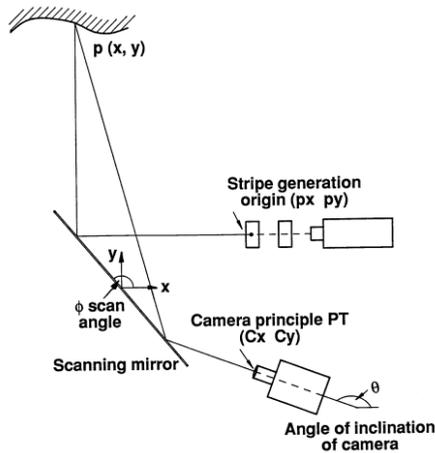


Abbildung 5. Aufbau eines Laserscanners [2]

und Beweglichkeit biete, wie Levoy in [11] erklärt. Somit war es möglich sogar Michelangelo´s Meißelspuren geometrisch zu rekonstruieren. Problematisch war in diesem Zusammenhang die(, von der Farbgebung her eigentlich sehr geeignete,)Marmoroberfläche der Statuen, da die kristalline Struktur des Marmors den Laser unter der Oberfläche streut und so zu Rauschen führt.

3.3 Methoden mit strukturiertem Licht (Structured Light)

Eine störungsunempfindliche Methode, die ohne teure Hardware auskommt ist die Verwendung von strukturiertem Licht als Erkennungshilfsmittel. In diesem Bereich wird sehr viel Forschung betrieben und es werden auch herausragende Ergebnisse erzielt, was mitunter an einigen Vorzügen dieser Methodenklasse liegt: Sie vereint die Stärken der beiden oben besprochenen Vorgehensweisen, ohne ihre Schwächen zu signifikant werden zu lassen. Einen hervorragenden Überblick über die verschiedenen Arten dieser Technik bietet [3].

Indem man das Objekt mit strukturiertem Licht beleuchtet, werden die korrespondierenden Bildpunkte leichter identifizierbar als bei der Photogrammetrie, was die algorithmische Komplexität und Störanfälligkeit vermindert. Ziel ist immer, korrespondierende Positionen auf den Referenzbildern mit eindeutigen Bezeichnern zu belegen. Hierfür gibt es nun unterschiedliche Ansätze.

Direktes Kodieren (Direct Codification) Das Prinzip: die jeweilige Stelle des Objekts wird mit einer Farbe oder Farbintensität(Phase) beleuchtet, welche direkt als Identifikator für die Zuordnung weiter verwendet werden kann. Die

Problematik dabei besteht darin, dass nur beschränkt viele, ausreichend unterscheidbare Codes zur Verfügung stehen.

In der einfachsten Variante [15] wird hierfür ein Schwarz-Weiß-Verlauf benutzt und beim Auslesen die jeweilige Graustufenintensität jedes Bildpunkts betrachtet. Daraus berechnet man wiederum die Verschiebung zum Referenzmuster und so die z-Koordinate. Dieses Verfahren ist sehr störungsanfällig und bedarf einer vorhergehenden Kalibrierung. Verbesserungen lassen sich erreichen, indem man jeden Farbkanal separat behandelt.

Wust und Capson [16] stellten 1991 eine Methode vor, die mit nur einem projizierten Muster auskommt: Sie kombinierten drei Sinus-Verlaufsmuster der Farben Rot, Grün und Blau, jeweils um 90° versetzt. Die Phasenverschiebung im aufgenommenen Bild kann dann durch die Formel

$$\psi(x, y) = \arctan\left(\frac{I_r - I_g}{I_g - I_b}\right)$$

berechnet werden, wobei $\psi(x, y)$ die Phase im Pixel (x, y) ist und I_r, I_g, I_b die jeweilige Intensität der Farbe Rot, Grün bzw. Blau. Diese Methode ist bereits für bewegliche Objekte geeignet, diese müssen allerdings relativ farbneutral sein und dürfen keine zu starken Einbuchtungen aufweisen. Wir kommen im Anwendungsbeispiel noch zu einer ähnlichen Methode, die diese Nachteile weniger signifikant aufweist.

Zeitbasiertes Multiplexing (Time Multiplexing) Diese Methode hat Ihren Namen daher, dass die Codewörter der einzelnen Bildpunkte sukzessiv durch mehrere hintereinander projizierte und abfotografierte Muster gebildet werden. Die Pioniere dieser Technik sind Potsdamer und Altschuler [17]. Sie encodierten erstmalig 2^m Streifen mit m Mustern. O.B.d.A. vertikale (horizontale funktionieren analog) schwarz-weiße Streifen werden auf das Objekt projiziert, wobei sich die Anzahl der Streifen mit jedem Muster verdoppeln. Die Anzahl der Streifen im letzten Muster stellt die (horizontale) Auflösung dar. In den abfotografierten Bildern werden nun alle Pixel identifiziert, die voll beleuchtet sind und bekommen eine 1 zugewiesen. Die restlichen erhalten eine 0. Das Bit, das ein Pixel vom ersten Muster erhält stellt das Most-Significant-Bit (MSB) dar, das vom letzten das Least-Significant-Bit (LSB), dazwischen analog. Daraus konkatenieren sich für jede Spalte eindeutige Codewörter und jedes Pixel einer Spalte weist am Ende das gleiche Codewort auf, was es einfach macht, die zum Referenzbild korrespondierenden Spalten zu identifizieren. Weniger störanfällig wird die Methode, wenn man die Reihenfolge jedes zweiten Spaltenpaares vertauscht, da dann benachbarte Codewörter eine Hammingdistanz von eins haben wie in Abb. 6 illustriert.

Mit dieser Technik kann die Kante einer Spalte mit Pixelgenauigkeit berechnet werden. Um noch akkuratere Ergebnisse zu erhalten änderten u.a. Rocchini et al. [8][9] das Schema insoweit ab, als dass sie Schwarz und Weiß gegen Rot und Blau austauschten und zwischen diesen beiden Farben eine ein-Pixel breite, grüne Linie einfügten. Ein Beispiel für eine Scanreihe mit dieser Mustervariante ist in Abb. 7 zu sehen. Die Position der Kanten des n-ten Bildes werden

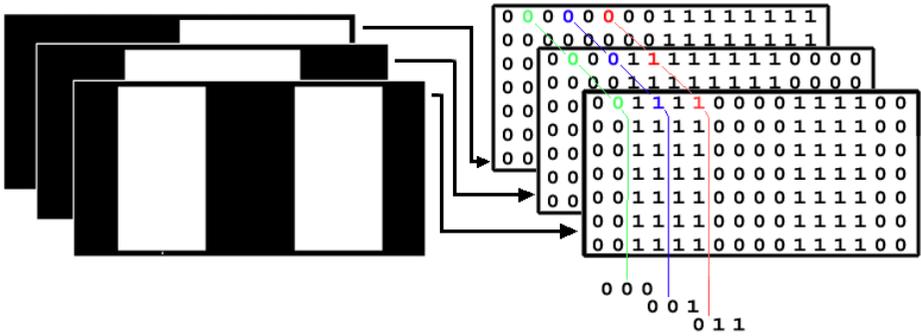


Abbildung 6. Schwarz-Weiß-Codierung mit einer Hamming-Distanz von eins

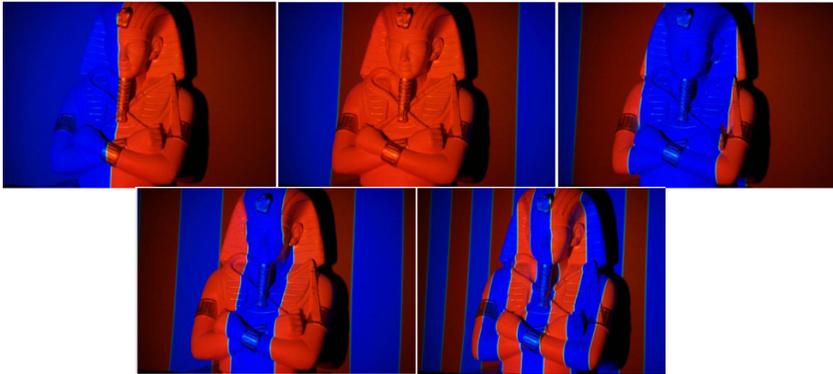


Abbildung 7. RGB-Muster [9]

dann auf Basis der Informationen errechnet, die man aus den vorhergehenden $n-1$ Bildern erhalten hat. Zuerst errechnet man die Position der grünen Linie durch den Übergang von Blau nach Rot und betrachtet dann die Funktion des grünen Farbverlaufs in den kritischen Pixeln. Die Stelle an der die 1. Ableitung der Intensitätskurve gleich Null ist, stellt die Position der grünen Linie mit Subpixelgenauigkeit dar.

Räumliche Beziehungsmuster(Spatial Neighborhood) Wie in der Photogrammetrie wird auch hier nach Ähnlichkeiten in den Pixeln der direkten Umgebung gesucht, allerdings mit Hilfe eines aufprojizierten Musters. Auch hier finden sich in der einschlägigen Literatur die verschiedensten Herangehensweisen. Vorgestellt werden soll an dieser Stelle die Methode von Ito und Ishii[18]. Diese benutzen ein Schachbrettmuster, wobei die Felder in einer aus drei Graustufen eingefärbt sind. Die Intensität des Grautons wird so gewählt, dass sie sich von der der vier direkten Nachbarfelder unterscheidet (siehe Abb. 8). Die Punkte, an denen vier Felder zusammenstoßen wird als Knoten bezeichnet und für

diese jeweils ein Codewort berechnet: Das Hauptcodewort eines Knotens (rot) errechnet sich aus den Intensitäten seiner angrenzenden Felder. Das Subcodewort (das eigentlich gesuchte Codewort, das den Knoten eindeutig identifiziert) ist die Kombination der Hauptcodewörter der benachbarten Knoten (blau) im Uhrzeigersinn. Es werden also für einen Knoten 12 umliegende Felder betrachtet (grün umrandet).

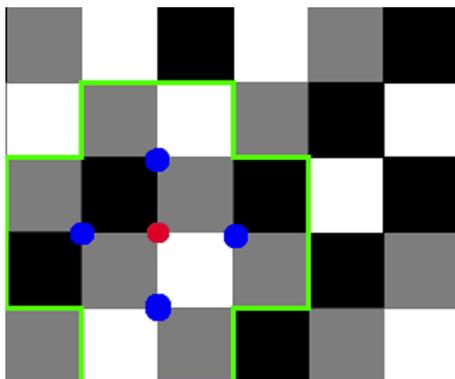


Abbildung 8. Schachbrettmuster nach Ito und Ishii

Anwendungsbeispiel: Echtzeitschnittstelle Hassebrook et al. beschreiben in [4] eine Möglichkeit auf Basis von strukturiertem Licht eine 3D-Echtzeitcomputerschnittstelle zu realisieren. Die Idee dabei ist zwei normalerweise sukzessive Vorgehensweisen so zu kombinieren, dass man mit einem Bild pro Akquisition auskommt. Verlaufsmuster mit Sinuscharakter (vgl. 3.3, Direktes Kodieren, [16]) wie in der „Phase Measurement Methode“ (PMP) verwendet, werden phasenverschoben und auf Streifenmuster mit unterschiedlichen spatialen Frequenzen aufmoduliert (siehe Abb. 9).

Die so enthaltenen Bilder werden zu einem einzigen Composite Pattern (CF) vereinigt. Das einzelne Bild, das man erhält wenn man das Objekt mit diesem Muster beleuchtet und abfotografiert, kann mittels Bandpassfilter wieder in einzelne Bilder zerlegt und mittels klassischer PMP-Decodierung ausgewertet werden. Realisiert man diesen Ansatz mit Projektionslicht im Infrarotbereich und einer Videokamera, hat man eine unsichtbare Computerschnittstelle die ein Mensch ohne Spezialausrüstung (Cyberhandschuhe, Tracker o.ä.) bedienen kann.

4 3D Benutzungsschnittstellen

Im allgemeinen ist die gängige Methode 3D-Inhalte zu erzeugen ein 3D-Modellier-Programm (wie z.B. [19]), bedient mit Tastatur und Maus. Es liegt nahe für Arbeiten an räumlichen Inhalten auch Werkzeuge zu benutzen, die in

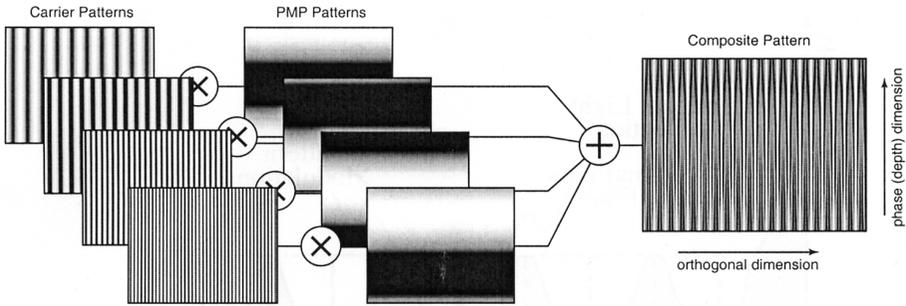


Abbildung 9. Generierung eines Composite Pattern [4]

Beziehung zum Raum stehen. Das folgende Kapitel zeigt einige Ansätze, den virtuellen mit dem realen Raum stärker zu verbinden.

4.1 Nicht-interaktive Schnittstellen

Vermittlung von realen Strukturdaten an den Computer (ohne direkte Rückkopplung während des Erzeugens der Struktur) ist die intuitivste und für den Benutzer einfachste Möglichkeit virtuelle Rauminhalte zu erzeugen. Die gewünschte Form wird manuell erstellt und der Schritt in die Virtualität erfolgt davon unabhängig. Ein Vorteil hier ist, dass man anhand von manuell definierten Vorlagen (Templates) die virtuelle Struktur mit zusätzlicher Logik ergänzen kann, was es dem Benutzer noch einfacher macht, sehr schnell relativ komplexe Konstruktionen zu erzeugen.

Manuell erstellte Modelle, automatisch animiert Reale Objekte aus verformbarem Material wie Ton oder Knetmasse herzustellen fällt oftmals leichter als erst die Benutzung einer komplizierten Computerschnittstelle zu erlernen. Anderson et al. nehmen in [7] die Herangehensweise von Kindern zum Vorbild, die einfache Modelle aus Knete bauen und die Details durch ihre Vorstellungskraft ergänzen. In diesem Sinne können einfache Objekte wie z.B. eine Figur, ein Auto, ein Haus, o.ä. geformt, hinterher gescannt und mit Textur, Details und Animation ergänzt werden (Abb 10). Dazu bestimmt man im Vorherigen eine Menge von Templates, die als Prototyp für die jeweilige Formenklasse dienen und ergänzt diese mit zusätzlicher Logik, wie z.B. wo sind bei einem Vierfüßler die Beine, wie bewegt er diese beim Laufen usw. Das gescannte Modell kann nun vom Computer mit den Templates verglichen werden und die Logik des Templates mit dem höchsten Grad an Übereinstimmungen wird dann auf das Modell übertragen (Abb. 11).

Der gleiche Ansatz kann auch auf Baumaterial angewendet werden, welches seine Struktur selbstständig an den Computer vermitteln kann. Anderson et al. statteten eine Art Legobaustein hierfür entsprechend mit Elektronik aus. Die Anschlußstifte (siehe Abb. 12) die aus einem inneren Kontakt und einem äußeren



Abbildung 10. Dreidimensionales Modellieren mit realen Materialien [7].

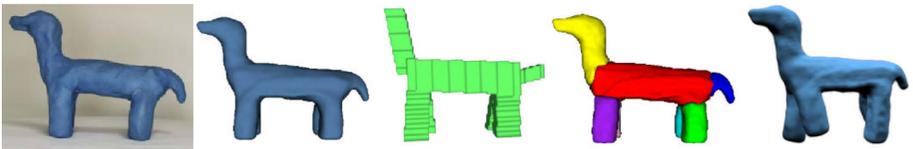


Abbildung 11. Das reale Objekt, Scan des Objekts, das passende Template aus einer Bibliothek, Match der vordefinierten Zonen am virtuellen Objekt, animiertes virtuelles Objekt [7].

Mantel bestehen, sind abwechselnd mit Phase und Erde am Mantel verdrahtet. So ist eine Stromversorgung sichergestellt, vorausgesetzt, ein Baustein ist nicht „unnatürlich“ angebaut (nur auf einem oder nur auf den diagonalen Konnektoren). Der innere Kontakt ist somit frei für die Kommunikation zwischen den Bausteinen und so zum Feststellen der Position jedes Blocks. Auch wurde zusätzliche Logik ergänzt, um das verarbeitende Programm selbstständig Wände, Dächer Fenster usw. erkennen zu lassen und diese auf verschiedene Arten mit passenden Texturen, Details und Formkorrekturen etc. zu versehen (siehe Abb. 13).

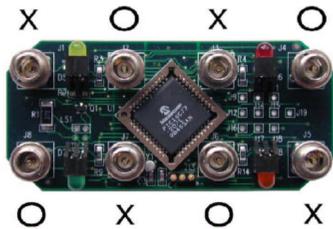


Abbildung 12. Die Konnektoren der Bausteine mit alternierender Belegung [7].



Abbildung 13. Intelligente Bausteine: Das reale Modell, die Darstellung der Strukturdaten im Computer, vom Computer berechnete Erscheinungsergänzung [7].

4.2 Interaktive Schnittstellen

Der eben besprochene Ansatz ist für viele Anwendungen adäquat und auch ausreichend, allerdings ist zumeist eine direkte Rückkoppelung wünschenswert, häufig auch nötig, z.B. beim kreativen, rein virtuellen Arbeiten wenn keine realen Objekte verwendet werden sollen oder können, und natürlich auch bei jeglicher Mensch-Maschine-Interaktion, bei der Funktionen in Echtzeit gesteuert werden sollen.

Twister - 3D-Modellierung im Raum Llamas et al. [12][13] haben ein Werkzeug zur Manipulation von virtuellen Gegenständen im realen Raum, genannt „Twister“ entwickelt. Hierfür hält der Benutzer in einer oder in beiden Händen je einen getrackten (vom Computer ständig lokalisierten) Würfel mit drei Knöpfen. Die Tracker unterstützen alle 6 Freiheitsgrade wodurch sämtliche Gesten der Hände des Benutzers erfasst werden können. Die Knöpfe erfüllen die folgenden Aufgaben: 1) Greifen. Wird der Knopf gedrückt, wird ein der grafischen Repräsentation des Würfels (Gizmo) naher Teil des virtuellen Objekts festgehalten und kann verformt werden. 2) Umgreifen. Da der Mensch durch seine Physiognomie Drehbewegungen maximal in einem Winkel ca. 180 Grad durchführen kann, erlaubt dieser Knopf, den augenblicklichen Zustand des Objekts einzufrieren, die Hände in eine angenehmere Position zu bringen um dann an der selben Stelle weiter zu arbeiten, mit der gleichen Auswahl an Modellknoten (Vertices). 3) Affektweite: „Wie fest“ die einzelnen Punkte des Objekts gehalten werden, wird durch eine mit diesem Knopf einstellbare Funktion festgelegt. Ähnlich der in 3Ds max 4 und 5 [19] verfügbaren Funktion der „weichen Auswahl“ von Vertices, bestimmt der Abstand eines Vertex vom Greifpunkt des Gizmo, wie stark sich die Manipulation auf diesen auswirkt (in Abb. 14 angezeigt durch die roten und grünen Bereiche). Um so weiter also ein Punkt vom Gizmo im Moment des Greifens entfernt ist, desto mehr schwindet die Manipulationsstärke. Die Geschwindigkeit des Schwindens ist mittels einer Kugel einstellbar, die auf Knopfdruck erscheint und ihren Radius durch ziehen verändert.

Die Implementation der zuletzt genannten Funktion verdient besondere Aufmerksamkeit. Benutzt man den Twister beidhändig und liegen die Greifpunkte so nahe beieinander, dass Sie sich gegenseitig beeinflussen würden, passt sich die Greifstärkefunktion an, um eine natürlichere Deformation zu erzielen. Die beiden Stärkesphären stauchen sich so, dass die beiden Greifpunkte nicht mehr

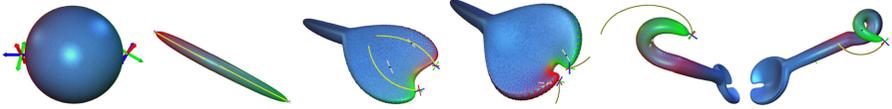


Abbildung 14. 3D Modellierung mit Twister [12].

im wechselseitigen Einflussbereich liegen (siehe Abb. 15).

Die verwendete Hardware: zwei Polhemus Fastrak trackers, NVIDIA Quadro4 900 XGL Grafikkarte auf einem Dual Pentium 3 866 Mhz mit 256 MB RAM. Bei bis zu 30.000 Vertices waren 20 fps möglich. Dabei wurde für eine weiche Darstellung die Anzahl der Polygonflächen herunter gerechnet. Auch möglich sei es, die Manipulationen erst auf einem Low-Polygon-Modell durchzuführen und aufzuzeichnen, um sie hinterher auf das eigentliche Modell anzuwenden.

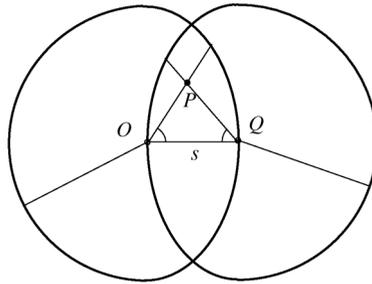


Abbildung 15. Stauchen der Einflußsphären beim beidhändigen Greifen an den Punkten O und Q [12].

Zange, Pistole und Lichtschwert als Werkzeuge zur Manipulation von Raumdaten Schkolne, Ishii und Schröder [6] experimentieren mit einem 3D-Interface, genannt „Silkworm“ das dem Benutzer erlaubt, virtuell DNA Moleküle im Raum zu erzeugen. Ausgabemedium ist hier eine Holobench (Tisch mit Oberfläche als Projektionsfläche für Projektoren kombiniert mit getrackter Shutterbrille, wodurch ein sehr immersiver räumlicher Eindruck entsteht), zur Eingabe werden drei Werkzeuge benutzt: Eine Zange (ähnlich einer Grillzange) zum greifen und bewegen der verfügbaren Einzelteile, ein Lichtschwert zum Zertrennen von Molekülverbindungen und eine Pistole. Mit dieser (siehe Abb. 16 rechts) werden Hydrogen- und Phosphatverbindungen zwischen Einzelteilen geschaffen. Die Analogie, die hierbei ausgenutzt wird, ist die der Klebepistole, um dem Benutzer ein vertrautes Werkzeug zu geben. Besonders innovativ an diesem Ansatz, der viel aus dem Bereich berührbare Benutzungsschnittstellen (Tangible Data) mitbringt, ist, dass die Vorteile von realer (Werkzeuge) und virtueller

Welt genutzt werden. Beispielsweise sind alle Werkzeuge, um zu viele Geräte (und damit auch zu viele störende Kabel usw.) zu vermeiden, virtuell mit Funktionen überladen. Die Zange (siehe Abb. 16 links), mit leichtem Druck um eine Base geschlossen, bewegt das ganze Molekül, mit starkem Druck bewegt man nur die gegriffene Base. Greift man abseits des Moleküls, kann man die ganze Szene verschieben, greift man mit beiden Zangen den leeren Raum, ist zusätzlich rotieren und skalieren möglich. Der Griff des Lichtschwerts kann umgeschaltet werden um als eine Art „DNA-Spender“ zu dienen. Die Funktionswechsel von Pistole und Lichtschwert werden durch Druck auf einen Menü-Knopf am Gerät ermöglicht. Direkt neben dem Werkzeug erscheint ein Auswahlmenü im Raum, so dass der Benutzer nicht durch Positionswechsel im Arbeitsfluss unterbrochen wird (siehe Abb. 17).

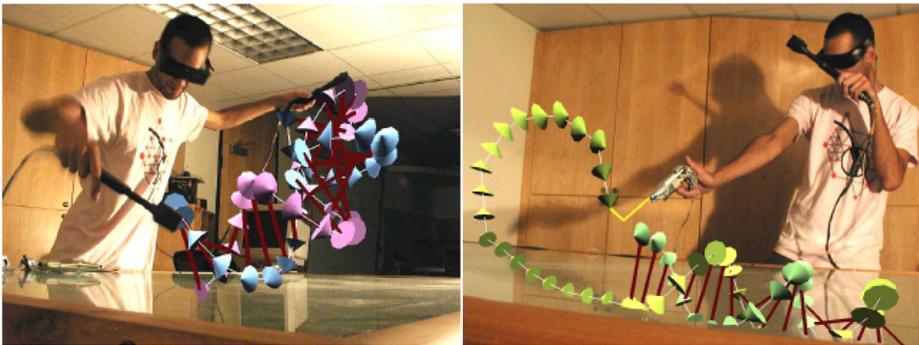


Abbildung 16. Silkworm: DNA - Modellierung an der Holobench [6] (Anm. d. Autors: Die Bilder sind mit Illustrationen ergänzt, da so nicht fotografierbar.)

Cavepainting - Malerei in 3D Hierbei handelt es sich um eine Benutzerschnittstelle für ein CAVE (CAVE Audio Visual Environment) womit der Benutzer nicht nur die Wände, sondern den gesamten Raum zum Malen zur Verfügung hat (siehe Abb. 18 links). Wie bei Silkworm (s.o) verwenden auch Keefe et al. [5] eine Kombination aus virtueller Datendarstellung und realen Werkzeugen zur Manipulation der Daten. Zum Auftragen der Farbe und Erzeugen von Formen wird ein getrackter Pinsel verwendet, der verschiedene Stricharten (auch plastische) realisieren kann. Um zwischen diesen zu wechseln, steht ein realer Tisch (siehe Abb. 18 rechts) als Kontrollelement bereit. Taucht man den Pinsel in einen der darauf platzierten Töpfe, wechselt die Strichart zu der, die durch Beschriftung dem jeweiligen Topf zugeordnet ist. Ein zusätzliches Kontrollelement stellt ein Cyberhandschuh für die nicht-dominante Hand dar. Zusammenführen von Daumen und Zeigefinger aktiviert einen Farbwahlwürfel, aus dem mit dem Pinsel die gewünschte Farbe selektiert werden kann. Schließt man Daumen und Mittelfinger, kann man die Größe des (virtuellen) Pinsels bestimmen, die als



Abbildung 17. Lichtschwert, Pistole und Zange, die Werkzeuge von „Silkworm“ mit Auswahlmenüs [6]

Icon an der Spitze des (realen) Pinsels angezeigt wird. Daumen und Ringfinger aktivieren den Translationsmodus, mit dem ein Objekt im Raum verschoben werden kann und der kleine Finger stellt den Skalierungsmodus zur Verfügung. Durch das Arbeiten im virtuellen Raum gelten einige Beschränkungen der realen Welt nicht. Zum Beispiel kann man mit einem Eimer virtuell Farbe nicht nur auf den Boden, sondern auf alle sechs zur Verfügung stehenden Flächen (Decke, Wände, Boden) tropfen lassen. Ein Zeittool kann dafür eingesetzt werden, Arbeitsschritte rückgängig zu machen oder das ganze Werk, gleich einem Film, in seiner chronologischen Entstehung zu betrachten.



Abbildung 18. Cavepainting [5]



Abbildung 19. Cavepainting - „Vor- und zurückspulen“ [5]

5 Fazit

3D Scanning, 3D Rekonstruktion und 3D Benutzungsschnittstellen eröffnen neue Möglichkeiten virtuelle Inhalte zu gestalten, mit ihnen die reale Welt zu ergänzen und mit dem Computer auf intuitive Art zu interagieren. Wir haben gesehen wie 3D Scanning durch Triangulation realisiert werden kann und wofür welche der verschiedenen Varianten angewandt werden können. Die Lasertriangulation ist zumeist teuer und langsam, allerdings für unbewegliche Objekte exzellent geeignet, weil akkurat und wenig fehleranfällig. Für lebende Objekte eignet sich die Photogrammetrie besser, da sie mit einer sehr kurzen Scanzeit auskommt. Unter Umständen kann diese Variante sehr preisgünstig sein. Da aber die Umgebungsbedingungen sehr gut kontrolliert werden müssen, kann das zusätzliche Kosten

aufwerfen. Eine Kompromisslösung ist die des strukturierten Lichts, wobei durch verschiedene Techniken ebenfalls kurze Scanzeiten erreicht werden können. Der algorithmische Aufwand ist niedriger und die Störeinflüsse leichter kontrollierbar. Allerdings erhöhen gerade bei Rundumscans mehrere Projektoren wieder die Kosten. Gerade auf diesem Gebiet lässt allerdings die derzeitige Forschung vermuten, dass eine 3D-Echtzeitcomputerschnittstelle ohne Hilfsmittel wie Cyberhandschuhe oder Trackern bald möglich wird.

Weiter wurden interaktive und nicht-interaktive Schnittstellen für die Gestaltung von dreidimensionalen virtuellen Welten vorgestellt und deren Vorzüge für verschiedenen Anwendungsgebiete diskutiert. Modelle, die aus physikalischen Baustoffen modelliert und anschließend gescannt werden oder Ihre Raumstruktur selbstständig an den Computer vermitteln können bieten einen sehr intuitiven Zugang zur 3D-Modellierung und die Möglichkeit, automatisch Ergänzungen und Verbesserungen vom Computer generieren zu lassen. Raumschnittstellen wie Twister, aber auch stark anwendungsspezifische Werkzeuge wie Silkworm und CavePainting, erlauben eine natürliche Manipulation von 3D-Daten, wie es im realen Raum ähnlich geschehen würde.

Literatur

1. P. Siebert, S. Marshall: Human body 3D imaging by speckle texture projection photogrammetry. *Sensor Review*, 20 (3), pp 218-226, 2000. <http://www.faraday.gla.ac.uk/publications.htm>
2. <http://www.faraday.gla.ac.uk/laser.htm>
3. J. Salvi, J. Pagés, and J. Batlle: Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827-849, 2004. http://eia.udg.es/~jpages/coded_light/
4. C. Guan, L. G. Hassebrook, and D. L. Lau: Composite structured light pattern for three-dimensional video. *Opt. Express* 11, 406-417 (2003). <http://128.163.147.208/lgh/research.html>
5. D.F. Keefe, D. A. Feliz, T. Moscovich, D. H. Laidlaw, J. L. LaViola Jr.: CavePainting: A fully immersive 3D artistic medium and interactive experience. *Symposium on Interactive 3D Graphics* (2001). <http://www.cs.brown.edu/~dfk/cavepainting/index.html>
6. S. Schkolone, H. Ishii, P. Schröder: Tangible + Virtual = A Flexible 3D Interface for Spatial Construction Applied to DNA. *CHI(2003)*. http://www.multires.caltech.edu/pubs/chi2003_schkolne_ishii_schroder.pdf
7. D. Anderson, J. L. Frankel, J. Marks, A. Agarwala, P. Beardsley, J. Hodgins, D. Leigh, K. Ryall, E. Sullivan, J. S. Yedidia: Tangible interaction + graphical interpretation: A new approach to 3D modeling. In *Proc. SIGGRAPH*, pages 393-402, New Orleans, LA, July 2000. ACM. <http://www.merl.com/publications/TR2000-013/>
8. C. Rocchini, P. Cignoni, C. Montani, P. Pingi, R. Scopigno, R. Fontana, L. Pezzati, M. Cygielman, R. Giachetti, G. Gori: 3D scanning the Minerva of Arezzo. In *ICHIM'2001*. <http://vcg.isti.cnr.it/publications/publications.htm>
9. C. Rocchini, Paulo Cignoni, C. Montani, P. Pingi, R. Scopigno: A low cost 3D scanner based on structured light. In A. Chalmers and T.-M. Rhyne, editors, *EG 2001 Proceedings*, volume 20(3), pages 299-308. Blackwell Publishing, 2001. <http://vcg.iei.pi.cnr.it/~rocchini/>

10. A. Marbs: Erfahrungen mit 3D Scanning am i3mainz.
11. M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Petreira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, D. Fulk: The digital michaelangelo project: 3D scanning of large statues. In Proc. ACM SIGGRAPH 2000. <http://graphics.stanford.edu/projects/mich/>
12. I. Llamas, A. Powell, J. Rossignac, C. Shaw: Bender: Deforming 3D Shapes by Bending and Twisting a Virtual Ribbon with both Hands <http://www.cc.gatech.edu/grads/1/Ignacio.Llamas/research.htm>
13. I. Llamas, B. Kim, J. Gargus, J. Rossignac, C. D. Shaw: Twister: A space-warp operator for the two-handed editing of 3d shapes. in Proceedings of ACM SIGGRAPH 03, Computer Graphics Proceedings, Annual Conference Series, 2003. <http://www.cc.gatech.edu/grads/1/Ignacio.Llamas/research.htm>
14. S.J. Marshall, G.T. Reid, S.J. Powell, J.F. Towers, P.J. Wells: Data capture techniques for 3-D facial imaging, Computer Vision and Image Processing, ed. A N Barrett, 248-275, Chapman and Hall (1991). http://www.faraday.gla.ac.uk/about3d_frame.htm
15. B. Carrhill, R. Hummel: Experiments with the intensity ratio depth sensor. in: Computer Vision, Graphics and Image Processing, Vol. 32 (1985), 337-358, Academic Press. <http://www.bobhummel.com/publications/listofpublications.htm>
16. C. Wust, D. W. Capson, Surface profile measurement using color fringe projection. Machine Vision and Applications 4 (1991) 193-203.
17. J. Potsdamer and M. Altschuler: Surface measurement by space-encoded projected beam system. in Computer Graphics and Image Processing, Vol. 18 (1982) 1-17
18. M. Ito, A. Ishii, A three-level checkerboard pattern (tcp) projection method for curved surface measurement. Pattern Recognition 28 (1) (1995) 27-40.
19. Discreet's 3ds max: <http://www4.discreet.com/3dsmax/>

Berührungsgesteuerte Benutzerschnittstellen

Benjamin Schlerf

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`benni.schlerf@web.de`

Zusammenfassung Diese Arbeit befasst sich mit berührungsgesteuerten Benutzerschnittstellen und aktuellen sowie zukünftigen Einsatzmöglichkeiten. Anhand des Phantom Haptic Device (PHANToM) wird das vom MIT entwickelte, wohl bekannteste und am vielseitigsten einsetzbare Ein- und Ausgabegerät vorgestellt. Bestehende Techniken der Mensch-Maschine-Interaktion aus den Bereichen Medizin und Virtual Reality werden aufgezeigt und anhand zum Teil futuristisch anmutender Experimente wie z.B. der transatlantischen Berührung, des MIT Touch Labs verdeutlicht.

1 Einleitung

Vieles von dem, was vor Jahren noch Stoff für Science-Fiction-Filme lieferte, ist durch die stetige Entwicklung im Bereich der Informatik bereits realisiert, oder zumindest greifbar nah. So können heute mittels modernster Technik Menschen über große Entfernungen miteinander in spürbaren Kontakt treten. Welche von all den Ideen in naher Zukunft alle realisiert sein werden, können wir heute natürlich noch nicht wissen. In möchte zunächst auf die aus heutiger Sicht relevanten Anwendungen und Perspektiven eingehen, die Dank der intensiven Forschungen auf diesem Gebiet ermöglicht wurden, und die für uns daraus resultierenden Möglichkeiten erläutern.

Bei berührungsgesteuerten Benutzerschnittstellen handelt es sich nicht etwa um einen Versuch, die etablierten Mensch-Maschine-Interaktionen durch Hightech unnötig komplizierter zu machen. Ganz im Gegenteil liegt meist der Gedanke zugrunde, der Mensch sollte sich im Umgang mit dem Computer genauso intuitiv und natürlich verhalten können wie er es in allen anderen Bereichen des Lebens tut. Wie auch Hiroshi Ishii und Brygg Ullmer [1] in der Einleitung zu ihrer These der greifbaren Bits schreiben, hat die Menschheit seit jeher komplizierte Berechnungen und genaue Messungen durchgeführt, und zwar schon lange vor der Erfindung des PCs. Dabei griffen unsere Vorfahren nicht auf ein Programm zurück und ließen sich per Knopfdruck ein Ergebnis ausgeben, sonder sie hantierten mit eigens für einen bestimmten Zweck erschaffenen Gerätschaften.

Seit der Einführung des PCs werden viele verschiedenen Aufgaben mit einem einzigen Gerät durchgeführt. Dies brachte zwar unwiderlegbare Vorteile, aber ebenso eine immense Abstraktion mit sich. Diese Abstraktion soll nun wieder vereinfacht werden,

indem komplizierte Vorgänge mit uns natürlich erscheinenden Bewegungsabläufen,, wie wir sie aus dem Alltag kennen, gekoppelt werden.

2 Das Phantom Haptic Device



Abb. 1. Phantom Haptic Device [13]

Das Erzeugen einer virtuellen Umgebung allein ist zwar schon ein großer Schritt und erleichtert uns die plastische Darstellung von komplizierten Szenarien und Modellen,, ist aber letztlich allein doch unbefriedigend. Erst dadurch, dass wir die Gegenstände in dieser virtuellen Realität auch berühren können, entsteht ein wirklich realistischer Eindruck. Man kann generell sagen, je mehr Sinne bei der Aufnahme von Reizen beteiligt sind, umso realistischer wirkt die Situation auf den Menschen, was ja das Ziel einer virtuellen Realität ist. Eine Person in eine vom Computer erzeugte Welt zu integrieren, in der sie sich so real als möglich bewegen und so natürlich wie möglich handeln kann.

Zu diesem Zweck baut und vertreibt die Firma SensAble Technologies [8] das am MIT entwickelte PHANToM. Es gehört in die Gruppe der Force-Feedback-Geräte, die bei Interaktion Kräfte an den Benutzer zurückgeben.

Wie (Abb. 1) zu sehen ist, ähnelt das PHANToM optisch einem Roboterarm mit mehreren Gelenken, um eine Bewegung in jede Richtung zu ermöglichen. An seinem Ende befindet sich ein Art Griffel, welchen der Benutzer in die Hand nimmt. Alternativ kann sich anstatt des Griffels auch eine art Fingerhut befinden, in den man einen Finger legt. Die Bewegungen des Griffels werden (Abb. 2) analog in die Bewegungen eines Gegenstandes in dem virtuellen Raum umgesetzt. Sobald dieser Gegenstand im Raum mit einem anderen Gegenstand kollidiert, gibt das PHANToM diesen Widerstand an den Benutzer zurück und simuliert somit einen Zusammenstoß bzw. eine Berührung.

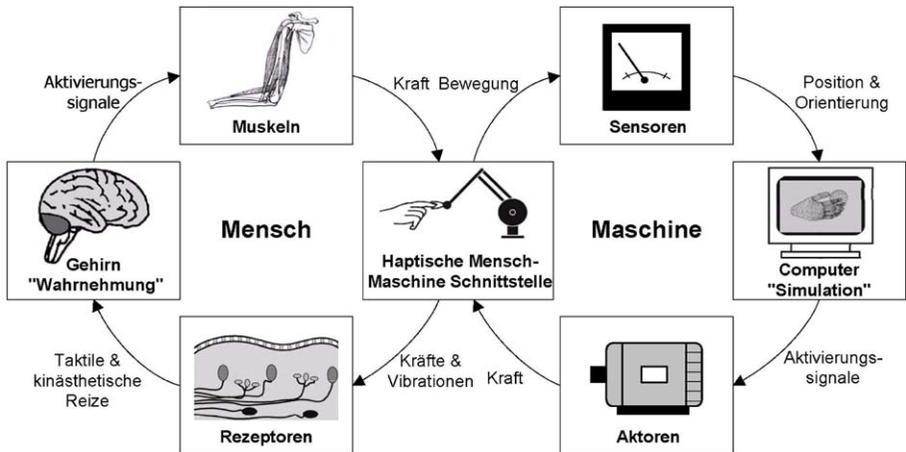


Abb. 2. Abläufe bei einer Kollision im Virtuellen Raum [6]

Um dem Nutzer einen realistischen Eindruck zu vermitteln, müssen drei Komponenten gegeben sein:

- Ein Modell das sie geographischen Gegebenheiten und die Materialeigenschaften bestimmt,
- Eine Berührungsschnittstelle die die Position des Nutzers verfolgen und die wirkenden Kräfte überträgt,
- Ein Algorithmus der die Kräfte in Abhängigkeit vom Verhalten des Benutzers berechnet.

2.1 Hardware

Der Einsatz des PHANTOMs ist sehr Rechenaufwendig. Die besten Ergebnisse werden daher erzielt, wenn das Gerät an einen eigenen Rechner angeschlossen ist. Die notwendigen Geometrieinformationen können ihm von einem Grafikrechner übermittelt werden. Eine andere Möglichkeit wäre ein Dual-CPU-PC der dank des zweiten Prozessors mehr Rechenleistung zur Verfügung stellt. In der Regel befindet sich das PHANTOM auf dem Desktop. Es kann jedoch auch, wie bei einem Versuch an der Iowa State University, auf einen Wagen montiert, und somit mobil genutzt werden.

Bei dem Versuch [3] ging es darum, inwiefern die Einbeziehung des Tastsinns sich auf die Wahrnehmung in einer virtuellen Umgebung auswirkt. Zu diesem Zweck wurde ein Raum gebaut, dessen vier Wände, die Decke und der Boden durch insgesamt sechs Bildschirme ersetzt wurden auf denen die virtuelle Umgebung durch Pro-

jektoren abgebildet wurde. In dem Raum, der insgesamt 10x10x10 Fuß maß, befanden sich eine Person und ein PHANToM.

Da der physikalische Arbeitsbereich des PHANToMs jedoch nur 19x27x37 cm betrug, musste man sich etwas überlegen um das PHANToM an jedem Ort innerhalb des Raumes nutzen zu können. Man montierte das Gerät auf einen höhenverstellbaren Wagen. Dieser Wagen lief auf vier fixierbaren Rollen, um ihn mobil und zugleich standfest zu machen. Um dem Wagen in der virtuellen Realität seine Position zuweisen zu können, wurde an einem der Beine ein elektromagnetisches Ortungsgerät befestigt. Um elektromagnetische Interferenzen zu vermeiden bestand der Wagen überwiegend aus PVC. Zusätzlich wurde eine spezielle Software entworfen, um die Differenz zwischen den Ausmaßen des physikalischen Arbeitsbereichs des PHANToMs und denen des virtuellen Raumes, so auszugleichen, dass das PHANToM schließlich in einem beliebig großen Raum benutzt werden konnte.

2.2 Software

In dem eben genannten Versuchsaufbau wurde die Bildschirm-basierte virtuelle Umgebung anhand der VRJuggler Software kontrolliert. Dabei handelt es sich um ein Programm, das für eine Vielzahl von Geräten aus dem Bereich der virtuellen Realität angewendet werden kann. Es stellt Bibliotheken und Routinen zur Verfügung um beispielsweise das Orten des Gerätes im virtuellen Raum zu ermöglichen.

Die Zusammenarbeit des VRJuggler und der PHANToM's Ghost Software, über die das PHANToM gesteuert wird, realisiert der PHANToM Treiber. Der Treiber ist ein C++ Klasse, die die Bewegungen umrechnet, den Ort des PHANToMs und die Graphik in der virtuellen Umgebung berechnet und mit sämtlichen anderen Anwendungen der Simulation kommuniziert.

Ein Grund, warum für die benötigten Programme so hohe Rechenleistung benötigt wird liegt darin, dass der Tastsinn des Menschen wesentlich besser ausgeprägt ist als zum Beispiel sein Sehsinn. Der Computer muss nur 20 Bilder pro Sekunde berechnen um unser Auge davon zu überzeugen, dass es sich nicht bloß um eine sequentielle Abfolge von Bildern sondern um eine homogene Bewegung handelt. Beim Tastsinn ist dies wesentlich aufwendiger da er um ein vielfaches sensibler ist. Erst ab einer Rate von 1000 Impulsen nehmen wir die übertragenen Kräfte als Bewegung wahr. Liegt die Rate unter dieser Anzahl, so wird der Reiz als unrealistisch eingestuft. Die Software muss also 1000-mal pro Sekunde den zu übermittelnden Kraftaufwand neu berechnen.

Besonders aufwendig ist auch die Kollisionserkennung. Für den Einsatz eines Force-Feedback-Gerätes ist diese natürlich besonders wichtig. Wird von der Software zu spät erkannt dass ein Gegenstand berührt wurde, kann dies dazu führen, dass der Benutzer in einen virtuellen Gegenstand eindringt. Abhängig davon, welches physika-

liche Modell bei der Programmierung der Software gewählt wurde, werden nun extrem hohe Kräfte berechnet. In der Regel steigen die Kräfte an der Oberfläche exponentiell an, was dann ein Herausschleudern aus dem Gegenstand bewirkt.

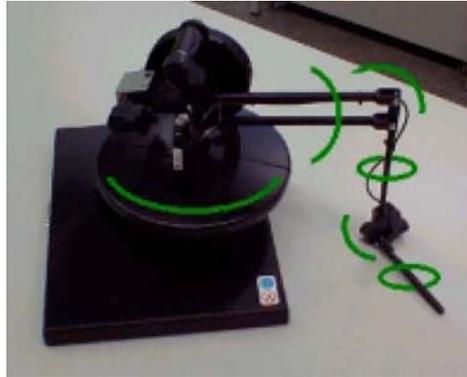


Abb. 3. Freiheitsgrade des PHANTOMs [7]

Dank der sechs Gelenke (Abb.3.) besitzt das PHANTOM sechs Freiheitsgrade. Die Position wird also durch ein Sechstupel beschrieben, das die Stellung jedes Gelenks beinhaltet. Um nun die Position in der virtuellen Umgebung zu erhalten, müssen diese Gelenkwinkel nun in kartesische Koordinaten transformiert werden, die anschließend weiter berechnet werden.

2.3 Einsatzgebiete

Der Einsatz von Force-Feedback-Geräten wie dem PHANTOM ist in vielen Bereichen sinnvoll. Bei der Erstellung oder Anwendung von behindertengerechter Software zum Beispiel. Oder in der Automobil- oder Flugzeugindustrie, wo die Entwicklung von Prototypen oft immense Kosten verursacht. Hier können Einzelteile vorab auf ihre Passgenauigkeit überprüft werden. Also direkte Operation von Objekten in CAD-Systemen. Auch in neueren Computerspielen, vor allem in Action-Spielen wird verstärkt auf den Einsatz von solchen Endgeräten gesetzt. Da die Industrie in diesem Zweig enorme Umsätze macht, scheint es auch besonders lukrativ dort weiter zu forschen. Ein anderes bereits weit entwickeltes Einsatzgebiet indem bereits große Fortschritte erzielt wurden ist die Medizin. Seien es Operationen an Patienten in fern abgelegenen Regionen, wie zum Beispiel bei Militärischen Einsätzen in abgelegenen Genden, auf Schiffen auf hoher See oder zur Integration von Frühwarnsystemen bei riskanten Operationen.

3 Medizin und Fernchirurgie

Viele Forschungseinheiten, vor allem in Ländern mit entlegenen Gegenden und weiten Flächen, befassen sich mit den Möglichkeiten der Telemedizin. Die Telemedizin findet Anwendung sowohl in der Pathologie als auch in der Dermatologie, da die Qualität der über das Internet übertragenen Bilder eine zuverlässige Diagnose zulässt. Die bemerkenswertesten, und für diese Arbeit am relevantesten Möglichkeiten ergeben sich jedoch in der Chirurgie. Im Jahr 2001, gelang es einem Chirurg, in New York (Abb. 4.), einer Patientin im 14 000 km entfernten Straßburg transatlantisch die Gallenblase zu entfernen.



Abb. 4. links: Arzt in New York, rechts: Patientin in Straßburg [14]

Das größte Problem bestand darin, dass die zeitliche Verzögerung zwischen Bild und Roboterarm nicht mehr als 330 Millisekunden betragen durfte, um zu gewährleisten dass der Arzt, sollte es während des Operationsverlaufes notwendig sein, rechtzeitig auf unvorhergesehene Ereignisse reagieren konnte.

Durch diese Operationsmethode könnten den Patienten in Zukunft etwaige teure und aufwändige Reisen zu weit entfernten Spezialisten, auch ins Ausland, erspart werden. Natürlich ist eine solche Operation auch immer mit einem Risiko verbunden. Sollten bei der Operation Komplikationen auftreten, und es müsste z.B. von einem laparoskopischen Eingriff auf eine offene Operation umgestiegen werden, so kann der Arzt nicht einfach das Operationsbesteck wechseln und möglichst schnell handeln. Es würde wesentlich mehr Zeit beanspruchen, länger als bei einer herkömmlichen Operation. Darum muss in der Regel eine Person vor Ort sein, die sich notfalls um solche Hilfestellung kümmert. Es ist aber wohl nur eine Frage der Zeit bis sich eine Lösung für dieses Problem findet.

Hardware

Erst durch die Verbreitung des Internet und eine schnelle Datenübertragung ist es möglich eine solche Zeitverzögerung in den Griff zu bekommen. Darin liegt aber

auch ein Problem. Oft existiert in abgelegenen Gegenden keine ausreichende Infrastruktur um derart große Datenmengen zu übermitteln.

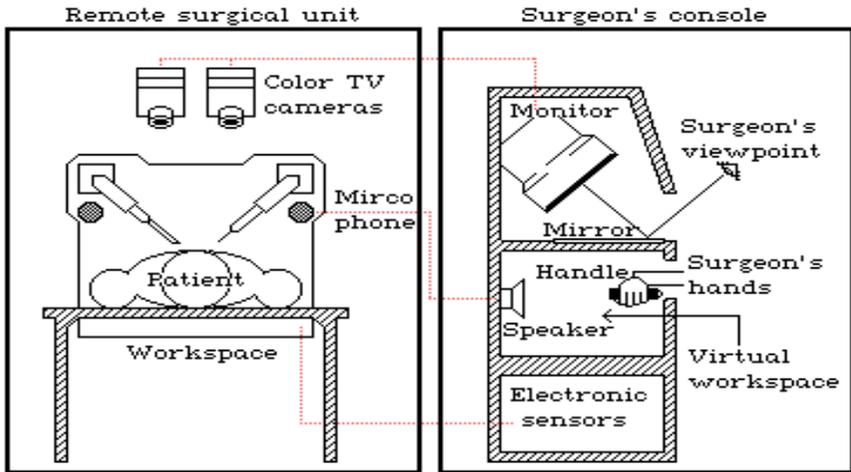


Abb. 5. Telechirurgischer Aufbau

Ist eine ausreichende Datenübertragung gewährt, so kann die Operation durchgeführt werden. Der technische Aufbau einer solchen Operation ist in (Abb. 5.) dargestellt. Während der Operation sieht der operierende Arzt auf einem Bildschirm die Gegend, in der operiert wird und die das Operationsbesteck das an einem oder mehreren Roboterarmen befestigt ist. In seiner Hand hält der Arzt Werkzeuge (Abb. 6), die dem echten Operationsbesteck nachempfunden sind (Abb. 7) und wobei jedes mit einem PHANToM verbunden ist. Die Kräfte, die durch den Widerstand beim durchtrennen mit dem Skalpell oder beim wegziehen beispielsweise mit einem Haken auftreten, werden auf das Besteck des Arzt übertragen.

Diese Methode wird auch zu Schulung von Ärzten genutzt. Dabei beschränkt sich das Einsatzgebiet nicht auf eine spezielle Ärztegruppe. An der Universität von Glasgow beispielsweise üben junge Veterinärmediziner an Simulationen von Pferden verschiedene Eingriffe.

Dabei ergeben sich dreierlei Vorteile:

- zum einen die Sicherheit, das gilt sowohl für Mensch als auch Tier. Komplizierte Operationen werden erst nach ausreichender Übung durchgeführt.
- Zweitens besteht die Möglichkeit verschiedenartige Eingriffe zu simulieren, auch solche, die in der Praxis eher selten auftreten, und daher oft falsch behandelt werden. Damit sinkt das Gefahrenrisiko eines fehlerhaften Eingriffs.

- Drittens können die Kosten gesenkt werden, da die Studenten in der Regel recht schnell damit umgehen und noch dazu beliebig oft üben können. Öfter als es an Lebenden Wesen möglich wäre.



Abb. 6 Simulation einer OP [16]

Für die Zukunft wäre auch die Realisierung eines Frühwarnsystems (Abb. 8.) denkbar. Der Computer müsste dazu in der Lage sein, auf eine Datenbank der menschlichen Anatomie zuzugreifen. Sobald der Operierende Arzt Gefahr läuft, wichtige Organe oder Arterien zu beschädigen würde er gewarnt, oder der Eingriff würde ihm im extremsten Falle sogar verweigert. Die letzte Entscheidung müsste dennoch der Arzt treffen. Schließlich gibt es immer wieder Ausnahmen, die besondere Behandlungsmethoden erfordern.

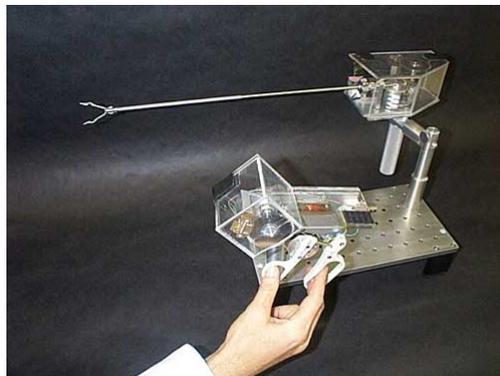


Abb.7. Besteck [15]



Abb. 8. mögliches Warnsystem während einer Operation [16]

4 Transatlantische Kooperation

Am 23. Mai 2002 gelang einer Forschungsgruppe des MIT in Zusammenarbeit mit einer Gruppe des University College London (UCL) die erste transatlantische (Abb. 9.) Berührung [2]. Nicht wie bei der Fern-Chirurgie waren an einem Ende ein Mensch und am Anderen eine Maschine, sondern auf beiden Kontinenten konnte ein Mensch die Bewegungen seines Gegenspielers spüren.

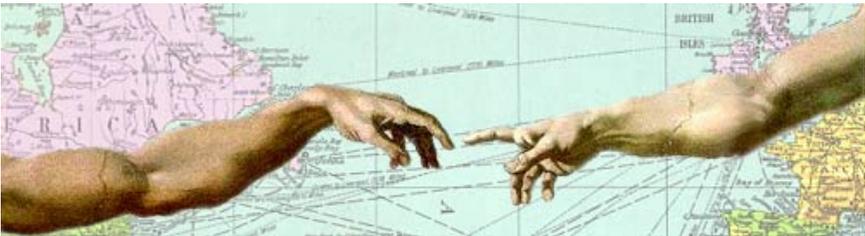


Abb. 9. transatlantische Berührung [2]

In London sowie in Cambridge saßen die Teilnehmer vor einem Bildschirm auf dem ein dreidimensionaler Raum zu sehen war. In der Mitte des Raumes (Abb. 10.) befand sich eine Kiste, links und rechts davon eine Markierung, die je einen der beiden Versuchsteilnehmer darstellte. Jeder der beiden Punkte wurde durch ein PHAN-ToM Haptic Device kontrolliert, dass sich neben beiden Bildschirm befand. Die beiden Punkte sind frei im Raum navigierbar.



Abb. 10. Blick auf den Bildschirm mit der Box [2]

Durch „drücken“ gegen die Kiste kann sie im Raum verschoben werden. Hierzu sind keine großen Kräfte notwendig, da die Koeffizienten so programmiert wurden, dass das Gewicht der Kiste 0.1 kg beträgt.

Um festzustellen inwieweit der haptische Einfluss die Koordination beeinflusst, wurden die 20 Versuchspersonen die in Cambridge an dem Experiment teilnahmen, in zwei Gruppen eingeteilt. Die eine Gruppe hatte nur die Bildschirmausgabe als Orientierung, die andere noch zusätzlich die Ausgabe des PHANTOMS.

Die beiden Computer waren über ein Internet2 Netzwerk verbunden, was Vorteile gegenüber einer normalen Internetverbindung hat, da eine schnellere Paketvermittlung möglich ist. Die durchschnittliche Verzögerung betrug zwischen 150 und 200 Millisekunden. Unser Körper braucht für die Übermittlung eines Reizes von der Hand zum Gehirn etwa 30 Millisekunden. Alle Bewegungen mussten also langsamer ausgeführt werden, da ansonsten eine als realistisch empfundene Synchronisierung nicht möglich gewesen wäre.

Mandayam A. Srinivasan, der Leiter der Gruppe des MIT meinte dazu „Even in our normal touch, there is a time delay between when you touch something and when those signals arrive in your brain. So, in a sense, the brain is teleoperating through the hand“ („Sogar bei unseren normalen Berührungen gibt es eine zeitliche Verzögerung zwischen dem Berühren von etwas, und dem Ankommen dieser Signale im Gehirn. Auf gewisse Weise operiert somit das Gehirn über unsere Hand fern.“).

Zuerst sollten sich die Versuchspersonen in einer kurzen Eingewöhnungsphase mit dem Phantom Haptic Device vertraut machen. Anschließend sollten sie gemeinsam mit der anderen Versuchsperson die Kiste anheben und so lange wie möglich in der Schwebe halten. Zuletzt füllte jeder Teilnehmer einen Fragebogen aus. Die Fragen drehten sich vor allem darum, ob während des Versuchs der Eindruck entstand, dass man mit einer anderen Person interagierte, nicht lediglich mit dem Computer. Falls ja, wie real dann dieser Eindruck empfunden wurde.

Die Auswertung ergab, dass alle Personen der Gruppe, die sowohl visuellen als auch haptischen Einflüssen ausgesetzt waren, die Erfahrung realistischer Einstufung als die Personen der anderen Gruppe ohne haptische Unterstützung. Ziel der Wissenschaftler ist nun eine Optimierung des Netzwerkes und der Verbindungen, um eine noch authentischere Umgebung zu schaffen.

5 Tangible/Greifbare Bits

Eine andere Forschungsgruppe des MIT, die Tangible Media Group, geht noch einen Schritt weiter. Bei seinen Forschungen geht es Professor Hiroshi Ishii, dem Gründer und Direktor der Gruppe, und seinen Mitarbeitern darum, die digitale mit der physischen Welt zu vereinen. Ein Zusammenspiel von Bits und Atomen.

Der Innovation ihrer Mensch-Maschine-Interaktion besteht in erster Linie darin, dass der Mensch sich nicht wie bisher üblich an seinen Computer setzt um dann mit der Tastatur Eingaben zu tätigen. Stattdessen versucht man, physisch greifbare Gegenstände so zu integrieren und mit Bits zu verknüpfen, dass man mit ihrer Hilfe Aufgaben erledigen und Informationen verarbeiten kann.

Die hierbei wichtigsten Konzepte sind

- Interaktive Oberflächen
- Das Verknüpfen von Bits mit Atomen
- Die Nutzung unserer umgebenden Medien wie Geräusche, Wasserbewegung etc.

Zwei Themen ihrer Forschungsarbeit werde ich im Folgenden vorstellen.

Das MetaDesk

Hierbei [10] lag die Idee zugrunde, dass man zwar Abschied nimmt von den klassischen Benutzeroberflächen (GUIs) die man von den Computer kennt, aber Elemente wie Fenster, Symbole oder Regler, die sich bereits weit verbreitet haben, durch reale Gegenstände verkörpert. Also das Erschaffen einer greifbaren Benutzeroberfläche (Tangible User Interface, kurz TUI).

Das MetaDesk besteht aus einer horizontalen, grafischen und projizierten Oberfläche, auf die die (Abb. 11.) abgebildeten Gegenstände gesetzt und bewegt werden können. Die Position dieser Gegenstände ist stets über elektro-magnetische Sensoren bestimmt. Ferner gibt es eine so genannte aktive Linse, die als Fenster fungiert. Sie entspricht einem tragbaren Display, und verkörpert ein Fenster durch das man Objekte noch genauer betrachten kann.

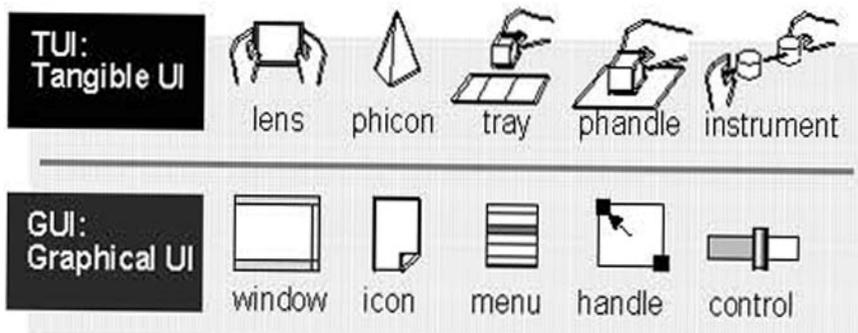


Abb. 11. GUI und TUI Element im direkten Vergleich [10]

Zum Testen wurde ein Programm entwickelt, welches einen Lageplan des MIT-Campus mit allen Gebäuden und der Umgebung beinhaltet. Um die Arbeit möglichst intuitiv und übersichtlich zu gestalten, wurden die Phicons (physikalisches Icon) echten Objekten nachempfunden. Wird zum Beispiel das MIT-Hauptgebäude nachempfundene Phicon auf das MetaDesk gelegt, so wird dort automatisch der Umgebungsplan des Campus angezeigt, und zwar so, dass das Hauptgebäude nun exakt auf der geographisch richtigen Stelle liegt. Dreht oder bewegt man das Phicon, so ändert sich auch die Lage des Stadtplans und passt sich an. Durch die aktive Linse, kann man sich nun das Hauptgebäude in dreidimensionaler Form aus verschiedenen Perspektiven ansehen. Auch zoomen ist möglich, indem man einfach ein zweites Phicon eines anderen Objekts auf dem MetaDesk aktiviert und den Abstand der beiden Symbole verringert oder vergrößert. Da der Abstand der beiden Objekte immer dasselbe Verhältnis hat, wird auf dem MetaDesk der Campus-Plan nun vergrößert bzw. verkleinert dargestellt.

Der AmbientRoom

Die Idee zur Weiterführung des MetaDesk basiert auf der Tatsache, dass der Mensch auf zweierlei Arten Informationen aufnimmt. Die Informationsaufnahme läuft über zwei Kanäle, den Primären und den Peripheren.

Zu den primären Informationen zählt man alles was mit den Händen getastet, mit den Augen gesehen oder direkt gehört wird, also alles worauf wir im jeweiligen Moment unseren Fokus legen. Mit peripheren oder hintergründigen Informationen sind die uns umgebenden Verhältnisse gemeint, wie Umwelteinflüsse oder jegliche Art von Nebengeräuschen. Alle diese Einflüsse finden ständig statt, ohne dass wir ihnen größeres Interesse beimessen. Erst wenn in unserer Umgebung etwas Außergewöhnliches unsere Aufmerksamkeit erregt, schwenken wir unseren Fokus darauf. Bei der Erschaffung des AmbientRoom (Abb.12.) ging es darum, sich ebendiese hintergründigen Einflüsse zunutzen zu machen.

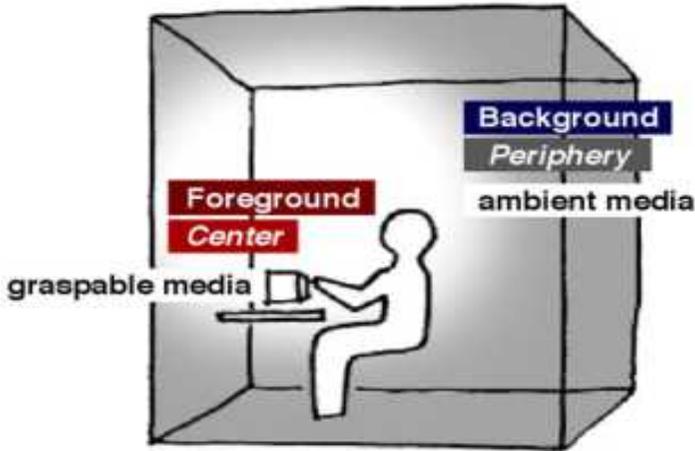


Abb. 12. Der AmbientRoom [10]

Durch den gesteuerten Einsatz von Licht, Wasser, Temperatur, Geräuschen etc, soll die hintergründige Aufnahme von Einflüssen die primäre Informationsverarbeitung unterstützen. Im wirklichen Leben können wir problemlos zwischen beiden Kanälen hin und her schalten. Zum Beispiel wenn beim Frühstück das Radio läuft und nebenbei noch Zeitung gelesen wird. Beginnt der Nachrichtensprecher mit den Neuigkeiten des Tages, werden die Informationen nicht mehr mit den Augen sondern mit den Ohren aufgenommen.

Man betrachte den Fall eines Spielzeug-Herstellers. Als Phicons nimmt man zum Beispiel die Modelle der neuesten Produkte. Legt man nun ein solches Phicon in eine dementsprechende Vorrichtung, so erscheint eine Projektion im Raum mit der Webseite des Produkts. Sanfte Töne, die über Lautsprecher im Hintergrund erzeugt werden, stünden beispielsweise für positive Verkaufszahlen, das Prasseln von Regen oder das Geräusch von Donner hingegen für schlechte Werte.

6 Diskussion

Es ist zu bedenken, dass, so präzise und intuitiv sie auch gewesen sein mögen, die Rechenmaschinen und Instrumente unserer Ahnen nicht über Nacht entstanden. Dieser Entwicklungsprozess dauerte Jahrhunderte. Setzt man dazu die erst sehr kurze Zeit, die auf dem Gebiet der berührungsgesteuerten Benutzerschnittstellen geforscht wird, in Relation, so kann man nur staunen über die Ergebnisse die in so kurzer Zeit erreicht wurden. Man ist dem Ziel der intuitiven, spürbaren Kommunikation von Mensch und Maschine ein großes Stück näher gekommen. Die Vorteile beider Medien zu vereinen, darin sehen Forschungseinheiten wie das Touch Lab ihre Aufgabe, und dass darin einiges Potential steckt, zeigt sich bereits heute. Neben den in dieser

Arbeit vorgestellten Experimenten und Einsatzbereichen gibt es noch eine Reihe anderer, kommerzieller Tätigkeitsfelder.

In der Industrie bei der Fertigung von Maschinenbauteilen wird diese Technik bereits eingesetzt. Auch in der Molekularbiologie, wo man unter dem Elektronenmikroskop Moleküle im Nanobereich nicht nur betrachten, sondern auch zerlegen und neu zusammenführen kann. Kunden könnten z.B. beim Autokauf die Möglichkeit wahrnehmen sich bei der Auswahl der Sitzbezüge durch die verschiedenen Varianten durchzutasten. Ganz abgesehen von sehbehinderten Menschen. Bisher bestanden für blinde Menschen nur sehr begrenzte Möglichkeiten virtuelle Welten zu betreten.

Die größten Schwierigkeiten sind schlicht technischer Natur, sei es die Zeitverzögerung über das Internet oder das Umsetzen von elektronischen Bausteinen entsprechend der Vorgaben. Auf allen diesen Gebieten wird geforscht und überall werden neue Fortschritte erzielt. Es ist wohl wie immer alles nur eine Frage der Zeit, den dem Erfindungsreichtum des Menschen sind kaum Grenzen gesetzt.

7 Schlussfolgerung

Über kurz oder lang werden die meisten dieser Entwicklungen wohl perfektioniert und vermarktet werden. Welcher Büroangestellte träumt nicht davon, anstatt stundenlang vor dem Computer zu sitzen und auf den Bildschirm zu starren und Eingaben nur mit der Maus oder der Tastatur zu tätigen, lieber spielerisch Informationen zu verarbeiten. Stress wird sich damit zwar nie ganz vermeiden lassen, und die Arbeit muss trotzdem erledigt werden, aber wenn diese sich in angenehmerer Weise und unserem naturell entsprechend verarbeiten ließe, Informationen unterbewusst wahrgenommen würden. wer wäre dann abgeneigt?

Literatur

1. Ishii H., Ulmer B. *Tangible Bits: Toward Seamless Interfaces between People, Bits and Atoms*, in Auszügen aus CHI'97, 234-241
2. Thomson E.A., *Researchers Reach Out To Touch Each Other*, MIT Tech Talk Ausgabe 47, Nummer 10, 30. Oktober 2002
<http://web.mit.edu/newsoffice/2002/touchlab-1030.html>
3. Fischer A., Vance J.M., *PHANTOM Haptic Device Implemented in a Projection Screen Virtual Environment*, Auszüge aus dem workshop on Virtual environments 2003
4. Ottensmeyer, *Cooperative Tele-Surgery*, 1995
<http://web.mit.edu/hmsl/markott/cooptelesurg.html>
5. Salisbury K., Brehm D., *Device Turns Computer into Tactile Tool, Lets User Touch Virtual Objects*, MIT Newsoffice, 11. März 1998
<http://web.mit.edu/newsoffice/nr/1998/phantom.html>
6. Thorsten Kühlen, *Virtual Reality Center Aachen, Virtuelle Realität in der Produktentwicklung*
7. Krämer B., *Haptische Ein- und Ausgabemedien in der Medizin*, Uni Karlsruhe
8. <http://www.sensable.com/>
9. J. Jordan, J. Mortensen, M. Oliveira and M. Slater, B. K. Tay, J. Kim, and M. A. Srinivasan, *Collaboration in a Mediated Haptic Environment*
http://www.cs.ucl.ac.uk/research/vr/Projects/Equator/papers/Documents2002/joel_presence_2002.pdf
10. Ishii H., Ulmer B. *The metaDesk: Models and Prototypes for Tangible User Interfaces*, in Auszügen aus UIST'97
11. Elizabeth Thomson, *MIT Touch Lab research spans basic science, virtual reality*, MIT News Office, 18. März 1999
<http://web.mit.edu/newsoffice/nr/1999/touchlab.html>
12. Brewster, Montgomery, Glendye, *Haptic Feedback in the Training of Veterinary Students*, Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow
13. Dowd P., *The Free form Modelling System*, 2002,
http://www.deskeng.com/articles/02/sept/spfeature/main_dig.htm
14. Brief Communications, *Nature* **413**, 379 - 380 27 September 2001,
http://www.nature.com/cgi-taf/DynaPage.taf?file=/nature/journal/v413/n6854/full/413379a0_r.html&filetype=&dynoptions=
15. Universität von Washington, *BioRobotics Laboratory, Surgical Technology*,
http://brl.ee.washington.edu/Research_Active/Surgery/Project_01/Project_01_P01.html
16. Forschungszentrum Karlsruhe in der Helmholtzgemeinschaft KISMET, *KISMET Medical Applications*, http://iregt1.iai.fzk.de/KISMET/kis_apps_med.html

Neue Benutzerschnittstellen für Spiele

Radostina Ruseva

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
ruseva@informatik.uni-muenchen.de

Zusammenfassung Entwicklung von Computerspielen ist ein wissenschaftliches Gebiet, das ständig neue Typen von Interaktionen zwischen Rechnern und Menschen erfordert. Die virtuellen Spiele sind intuitiver und konsistenter. Das erhöht die Benutzerfreundlichkeit und vermeidet die Verärgerung, die mit dem häufigen Wechsel zwischen dem physischen Echtzeitspiel und dem Rechner entsteht. Die folgende Arbeit gibt einen Überblick über Prototypen von neuen Benutzungsschnittstellen von Computerspielen. Der Schwerpunkt liegt dabei bei Interaktionsformen, die ohne zusätzliche Geräte am Benutzer auskommt. Es werden zuerst grundlegende Technologien zusammengefasst und erklärt. Die neuen Interaktionsmöglichkeiten, die sie anbieten, werden am Beispiel einiger Spiele behandelt.

1 Einführung

Die folgende Arbeit gibt einen Überblick über die neueste Entwicklung der aktuellen Benutzerschnittstellen für Computerspiele. Hierbei werden schon bekannte Schnittstellen aus dem Bereich Virtual Reality wie Datenhandschuh und –Brille ausgeschlossen. Der Schwerpunkt liegt auf Interaktionsformen, die ohne zusätzliche Geräte am Benutzer auskommt, da somit Spiele entwickelt werden, die intuitiv und konsistent sind - deswegen auch benutzerfreundlich und leichter zu erlernen. RF-ID Tags und intelligente Umgebungen (smart Roomware [05][25]) haben sich als die passenden Technologien durchgesetzt, um die physische mit der virtuellen Welt zu verkuppeln. In dieser Arbeit werden deswegen hauptsächlich Artikel über diese beiden Innovationen studiert. Am Anfang wird ein kurzer Überblick über die wichtigsten Technologien und ihre Vorteile gegeben. Ihre Umsetzungsmöglichkeiten in formen neuer Interaktionsmöglichkeiten werden dargestellt. Am Ende werden Prototyp-Spiele zusammengefasst, die im Rahmen verschiedener Projekte auf der Basis dieser Technologien entwickelt worden sind.

2 Grundlegende Technologien für neue Spiele

Die Notwendigkeit an neuen Technologien für Spiele wächst mit den steigenden Anforderungen an sie. Leichte Erlernbarkeit und intuitive Benutzbarkeit setzen voraus, dass der Nutzer (Spieler) unbeschränkt im Raum bewegbar ist. Es muss nach Möglichkeit nicht zwischen verschiedenen Schnittstellen manuell umgeschaltet werden. Die Konsistenz der Spiele soll durch eine leicht bedienbare und für den Nutzer logische Oberfläche gesichert werden.

2.1 Berührungsempfindliche Bildschirmoberflächen

Die traditionellen Brettspiele und andere "Tabletop"-Spiele sind populär wegen ihrer leicht zu bedienenden und intuitiven Interaktionsformen. Dabei ist die Kommunikation zwischen den Spielern ein wichtiger Gesichtspunkt, da es sich um eine Situation handelt, in der die Benutzer miteinander und mit dem Spiel interagieren, wie z.B. beim Kartenspielen. Neue Hardware Technologien erlauben mehreren Nutzern die gleichzeitige Bedienung einer Benutzeroberfläche, mit der sie einzeln interaktiv kommunizieren. Solche Tabletop Lösungen sind z.B. SMART DviT [09][11-16][27] und MERL DiamondTouch [09][26]. Sie ermöglichen das Zusammenwachsen der Vorteile von traditionellen Bettspielen und Computerspielen.



Abb.1 SMART DviT Oberfläche [17]

2.1.1 SMART DviT

SMART Digital Vision Touch ist eine berührungsempfindliche Bildschirmoberfläche (wie auf Abb. 1). In jedem Eckpunkt davon sind digitale Kameras eingebaut, die die Position von Fingern, Stiften oder anderen Objekten auf dem Bildschirm registrieren und weiter an einem digitalen Prozessor zur Verarbeitung übergeben. Die Technologie ist also nicht in der Oberfläche selbst eingebettet, sondern im Rahmen davon. [27] Das Bild davon unterscheidet sich mit ihrer hohen Auflösung, die zur perfekten Darstellung, hohe Geschwindigkeit und exaktes Messen und Wiedergeben von Interaktionen dazubringt. Die Oberfläche von SMART DviT ist „ungekünstelt“ (aus engl. „unaffected surface“), d.h. für die Interaktion mit dem Bildschirm werden keine speziellen Stifte oder andere technische Mittel gebraucht. Man kann sie auch mit den Fingern bedienen.

2.1.2 MERL DiamondTouch

MERL DiamondTouch benutzt eine hohe Anzahl an Antennen, die in der Oberfläche eingebaut sind, um die Koordinaten des Spielenden einzulesen. Sie sind in einer Art Matrix von Zeilen und Spalten geordnet, sodass jede Antenne ein einzelnes Signal empfängt und sendet. Diese Technologie ermöglicht auch die Interaktion mehrerer Spieler (bis zu 4) mit der Oberfläche (wie auf Abb. 2). Um Benutzer unterscheiden zu können, wird für jeden Nutzer ebenfalls eine Antenne benötigt. Diese wird normalerweise in seinem Stuhl eingebaut. [09][26]

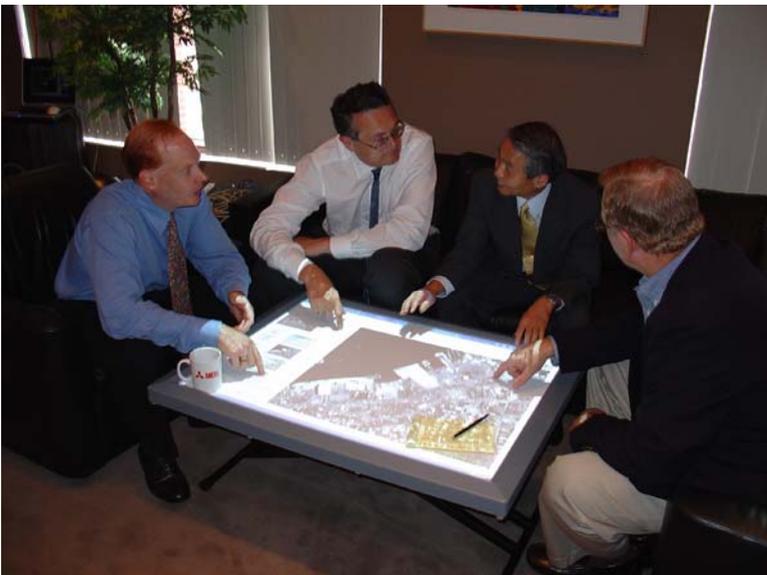


Abb. 2 MERL DiamondTouch ermöglicht die gleichzeitige Interaktion von bis zu vier Nutzern [18]

2.2 RFID Technologie

Eine andere Möglichkeit mit dem Rechner intuitiv zu interagieren, ist RFID-Tags für die Ermittlung von Position und Handlung auf dem Spielbrett zu benutzen. Dabei handelt es sich um berührungslose und Sichtkontakt unbedürftige Kommunikation. Ein RFID-System besteht aus Transpondern und Schreib-/Leseeinheiten mit Antenne, die auch mobil sein können. Die Entfernung, von der man RF-ID Tags auslesen kann liegt zur Zeit bis zu 30m, für Spiele ist aber der Nahbereich (5-50cm) relevant. Somit werden in der Eingabeoberfläche Radiofrequenzantennen eingebaut, die die Signalen, von in den Spielteilen selbst (wie auf Abb. 3) eingebetteten RFID-Transpondern, empfangen. Die letzten beinhalten alle Daten. Somit initialisiert oder beendet man ganz einfach das Spiel, indem man die Spielteile auf dem Brett drauflegt oder sie von ihm wegnimmt. [10][19]



Abb. 3 Mit RF-ID Tags versehenen Puzzle Teile [10]

2.3 Sensoren

Anderes bekanntes Forschungsfeld ist die Interaktion durch Sensoren. Die Benutzerschnittstelle sollte dabei möglichst kabellos sein. Die gleichzeitige Registrierung mehrerer Nutzer ist jedoch schwer umzusetzen und man braucht im Normalfall ein Vermittlungswerkzeug, damit die Kommunikation stattfinden kann.

2.4 Technologien für Draußen

Die oben beschriebenen Technologien mögen zwar eine intuitive Bedienbarkeit von Spielen sichern, sind auf die Bedingung in unmittelbarer Nähe beschränkt. Deswegen wird auch parallel an Technologien gearbeitet, die für Spiele im Freien geeignet sind.

2.4.1 GPS

Global Positioning System garantiert die exakte Positionierung mobiler Spieler, die räumlich voneinander unabhängig und beliebig entfernt sind. Die Koordinaten werden automatisch durch Sensordaten generiert. Das System stützt sich auf einen Satelliten, der mit einer ziemlich hohen Genauigkeit (bis zu 10m) schnell den Standort der Teilnehmer des Netzes bestimmen kann, dessen Koordinaten in einem GPS-Empfänger verarbeitet werden. Es ist allerdings ein vielversprechender technologischer Ansatz, der aber sehr fehleranfällig ist wegen der Abhängigkeit vom Wetter- und Landschaftsgegebenheiten ist. [08][28]

2.4.2 WLAN

Diese schon wohl bekannte kabellose Technologie ist schnell zu installieren, deckt eine große Fläche ab und funktioniert dann kostengünstig, was zu ihrer Beliebtheit gebracht hat. Das System besteht aus einem Accesspoint (der Sender) und WLAN-Karten (Empfänger) auf der Seite der Nutzer (siehe Abb. 4). Für die Spiele eignet sich am besten der 802.11 Standard und hauptsächlich der 802.11b mit 2,4 GHz Band und eine Geschwindigkeit von 11 Mbit/s. Besonders wichtiger Vorteil davon ist die hohe Reichweite (von bis zu 300m). Dieser Standard ist leider störungsanfällig, da auch andere Geräte auf ihm arbeiten. [08][29]



Abb. 4 WLAN Technologie ermöglicht die Mobilität der Spieler

2.5 Handheld und Mobile Computer

Zur Umsetzung von traditionellen Brettspielen in interaktive Tabletop-Spiele kann man sich im allgemeinem Fall mit einer gemeinsamen Benutzerschnittstelle für alle Spieler zufrieden stellen. Das alle Spieler gemeinsam mit der Spieloberfläche interagieren, ist sogar erwünscht. Für Rollenspiele oder andere Spiele, wo die Nutzer unter Konkurrenzbedingungen gegeneinander auftreten aber, gibt es oft eine Menge Informationen, die von den Gegnern nicht wahrgenommen werden dürfen. Unter solchen Bedingungen werden zusätzliche Benutzerschnittstellen, speziell für diese Informationen, benötigt. Als solche eignen sich besonders kleine mobile Rechner, Handhelds und jegliche andere Mobilgeräte.

2.5.1 Handhelds/PDAs

Sie lassen sich bequem in einer Hand halten, sind schnurlos und somit schränken die Mobilität des Nutzers nicht ein. Der Nutzer interagiert damit hauptsächlich durch das Farbdisplay, das selber berührungsempfindlich ist (siehe Abb. 5). Ihre Funktionalität ist zwar begrenzt, aber sie spielen ja eine unterstützende Rolle. Über Datenaustausch durch "Bluetooth" oder Infrarot können die Inhalte mit dem Rechner spontan synchronisiert werden. [30]



Abb. 5 Auf Handhelds wiedergibt man Informationen, die von anderen Spielern versteckt werden müssen [22]

2.5.2 UMTS Handy

Die kleinsten Multimediarechner (siehe Abb. 6), die unter allem auch nebenbei Telefonieren als Funktionalität anbieten, werden mittlerweile als Standard durchgesetzt. Sie lassen sich als Zusatzschnittstelle für neue Spiele einsetzen, da sie Klang und Video in Echtzeit übertragen können. [30]



Abb. 6 UMTS-Handys verfügen über die nötigen Funktionalitäten, um in neue Spiele genutzt werden zu können [21]

3 Interaktion

Die Entwicklung neuer Spiele setzt die Entwicklung neuer Benutzungsschnittstellen voraus. Traditionelle Eingabeschnittstellen wie Maus und Tastatur bieten keine ausreichenden Möglichkeiten zur gleichzeitigen kooperativen Interaktion zwischen mehreren Nutzern. Sie schränken außerdem die physische Interaktion auf die Eingabe durch die Finger ein und ist somit nicht für alle Nutzer geeignet.

3.1 Neue Benutzungsschnittstellen

3.1.1 Sensing GamePad

Durch Ausmessen des elektrostatischen Potentials des Körpers eines Spielers werden seine Fußbewegungen registriert, ohne dass dafür zusätzliche Sensoren gebraucht werden. Dafür werden in ein herkömmliches GamePad zwei Elektrode eingebaut

(siehe Abb. 7) und zu einem Hochwiderstandstromspannungs-Verstärker verbunden. Der Wechsel des elektrostatischen Potentials im Körper wird von einem Mikroprozessor registriert und dem Rechner durch USB Serienschnittstelle weitergeleitet. [02]

Abb. 7 Sensing GamePad misst elektrostatisches Potential durch Registrierung der Fußbewegungen des Spielers[02]



3.1.2 Ely the Explorer

Ely ist ein Spielzeug für Kinder im Alter von 6 bis 7 Jahre, entwickelt für interaktives Spiel unter mehreren Kinder, z.B. in der Schule, oder auch für individuelle Spiele (siehe Abb. 8). Sie "reisen" durch die Welt und entdecken fremde Kulturen. Elys können Ansichtskarten verschicken, Musik und Videos den Kindern am Display vorspielen. Sie dienen als Vermittlungsobjekt zwischen der Kinderphantasie und der neuen Technologie, mit dem Zweck ihre Kreativität und Wissen zu entwickeln. [01]



Abb. 8 Ely The Explorer ist ein interaktives Lernspielzeug, womit Kinder im Alter von 6/7 Jahren neue Kulturen entdecken können [23]

3.1.3 SenToy



SenToy ist ein Spielzeug (wie auf Abb.9), das als Eingabeschnittstelle für Emotionen speziell fürs Spiel FantasyA entwickelt worden ist. Es ist schnurlos und funktioniert durch die Sensoren, die in seinen Gliedern eingebettet sind. Die verschiedenen Positionen seiner Hände und Beine entsprechen den sechs grundlegenden Emotionen: Angst, Wut, Freude, Trauer, Überraschung und Schadenfreude. [03]

Abb. 9 SenToy ist eine Eingabeschnittstelle für Emotionen [4]

3.1.4 Roomware®

Unter Roomware® versteht man herkömmliche Wohnungsteile, die außer ihre alltäglichen Funktionen durch integrierte Kommunikationstechnologien auch multimediale Interaktionsmöglichkeiten anbieten. Das sind z.B. DynaWall® (interaktive berührungsempfindliche Wand), InteracTable® wie auf Abb.10 (vertikale berührungsempfindliche Projektionsoberfläche), CommChair® usw. Sie alle eignen sich hervorragend als Benutzerschnittstellen für neue Spiele, da sie deren Anforderungen (Mobilität, Konsistenz und Interaktivität) entsprechen. [05]



Abb.10 Roomware® Komponenten bieten Interaktionsdienste, ohne auf ihre alltägliche Funktionalität zu verzichten [24]

3.2 Neue Formen der Interaktion

Die Anforderungen an Benutzerfreundlichkeit und Akzeptanz für die Benutzerschnittstellen für neue Spielen dürfen wegen der neuen Formen von Interaktionen mit dem Rechner nicht vernachlässigt werden. Das Spiel soll aktiv die visuelle Welt aufnehmen können und darauf auch reagieren. Menschen kommunizieren nicht nur durch Sprache und Bewegungen untereinander. Sie haben noch Gefühle, sammeln mit der Zeit Erfahrungshintergrund und werden von der Umwelt beeinflusst. Die neuen Formen von Interaktion streben danach, diese Interaktionsmöglichkeiten mit dem Rechner zu entwickeln. Am Beispiel vom Ely The Explorer werden ins Spiel Emotionen eingelesen. Eine neue Form der Interaktion stellen auch die berührungsempfindlichen Spieloberflächen oder die vom elektrostatischen Potential abhängige SensPad. Ihre Umsetzung hat nichts künstlicher Intelligenz zu verdanken, sondern Algorithmen, auf denen herkömmliches Hard- und Software eingesetzt und implementiert werden.

4 Spiele-Plattformen

Eine Spielplattform ist eine Verkopplung von Hard- und Software, das die Umsetzung verschiedener Spiele erleichtert, indem es den Programmierern eine passende Schnittstelle anbietet. Im folgenden werden die Technologien (Bauteile, Implementierung und Design) für die Systeme STARS [05] [07] und TvIEWS [06] kurz zusammengefasst, da sie am besten den Anforderungen der zukünftigen Tabletop-Spiele entsprechen.

4.1 Hardware

4.1.1 Spieloberfläche

Als passende Spieloberfläche besonders bei Brettspielen eignet sich z.B. der InteracTable® [05]. Die von einem LCD Projektor erstellte Darstellung darauf ist dynamisch anzupassen. Es ist mittlerweile eine Auflösung von 1024x768px erreichbar. Der berührungsempfindliche Plasmabildschirm hat eine Größe von 70 zu 125cm und ist auf einer Höhe von 90cm gestellt. Über dem Tisch registriert eine Kamera die Positionen der Spielfiguren auf der Oberfläche. Darunter sind RF-ID Antennen eingebettet, um die RF-ID Tags Signale auf dem Tisch empfangen können. Der MERL Diamond-Touch Tisch erreicht eine Diagonale der Oberfläche von 107cm (4:3 aspect ratio). [09]

4.1.2 Vertikaler Bildschirm

Information, die von mehreren Spielern wahrzunehmen ist, lässt sich am besten auf einem großen vertikalen Display darstellen. Im Falle der DynaWall® handelt es sich um einen elektronischen Whiteboard mit Darstellungsbreite von 4,5m und –Höhe von 1,1m. Somit wird eine Auflösung von 3072x768px erreicht. [05]

4.1.3 PDA

Während die großen Displays für allgemeine Informationen, die alle Spieler betrifft, perfekt geeignet sind, werden für kleinere Informationsmengen, die aber persönlich Spielern zuzuordnen sind, andere Darstellungsmöglichkeiten gesucht. Die PDAs, die meistens in neuen Spielen eingesetzt werden stammen von Compaq oder Toshiba. Es handelt sich normalerweise um ein Standard PDA-Gerät (wie z.B. IPAQ 3950) mit eingebautem 802.11b WLAN, um die Verbindung zum Spiel zu ermöglichen.

4.1.4 Audio Devices

Lautsprecher sorgen für den öffentlichen Ton. Für jeden Spieler spezifische Privatnachrichten können z.B. durch integrierte Headsets wiedergegeben werden. Plattformen wie STARS bieten zusätzlich die Möglichkeit zur Spracheingabe und Sprachgenerierung basiert auf dem Microsoft Speech API. [05][07]

4.2 Software

Die neuen Benutzerschnittstellen für Spiele werden für 2D- sowie für 3D-Anwendungen angepasst. Die 3D-Spiele, wie z.B. PhantasyA, werden als 3D Entwicklungsumgebung implementiert, die auf OpenGL läuft [05][07]. Es wird der A* Algorithmus benutzt, um Emotionen zu erkennen [05][06][07]. Für andere Spiele, für deren Zwecke eine 2D-Darstellung ausreichend ist (wie z.B. bei Brettspielen), werden in Java (Java API) [06] programmiert.

4.3 Systeme

4.3.1 STARS

STARS ist eine Software Plattform, speziell für neue Brettspiele entwickelt, die auf Roomware Komponenten als Hardware aufbaut. Die ermöglicht die Integration verschiedener Geräte mit einem Spieltisch. Die Spiele selbst laufen auf einer Software Schicht, die das Programmieren von Spielen erleichtert, indem die Anzahl und die Art von Anschlussgeräten für das jeweilige Spiel nicht von Bedeutung werden. [05][07]

4.3.2 Tviews

Tviews ist eine interaktive Plattform zum Erfinden und Wiedergeben von Geschichten. Es ermöglicht die Spieler, ihre eigenen Multimedia Geschichten zu erfinden und anderen Nutzern, sie mitzuteilen. Sie benutzt das elektromagnetische Feld von Objekten, um ihre Position auf dem Brettspiel zu erkennen. Die Anwendung besteht aus drei Entwicklungsmodulen, die in Java implementiert sind: Positionregistrierung-, Wiedergabe- und "ErzählWeiter"-Modul. Das erste benutzt den Polling Mechanismus, das zweite die Apples QuickTime for Java API. [06]

5 Spiele

Im folgenden stelle ich kurz einige Beispiele für Spiele, die die oben beschriebenen Technologien benutzen und somit ihre Möglichkeiten kurz darstellen.

5.1 Ely The Explorer

Ely The Explorer ist gedacht als ein Lernspiel, das Kindern zwischen 6 und 7 Jahren das geografische, geschichtliche und kulturelle Wissen erweitern soll. Es kann von bis zu 3 Kindern auf einmal gesteuert werden. Jedes Kind verfügt über ein Ely, seine

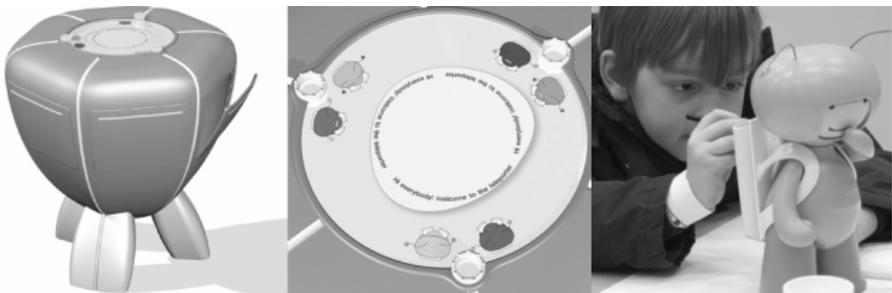


Abb.11 Der Teleporter (links), seine Benutzeroberfläche (Mitte) und ein Ely (rechts) [01]

Benutzungsschnittstelle. Es ist sowie eine Puppe, als auch eine virtuelle Figur, die am Bildschirm des Teleporters erscheint. Sie verfügen über einen Reiserucksack (wie auf

Abb. 11), der ein kleiner Display und eine Kamera eingebaut hat. Der Teleporter selbst hat eine berührungsempfindliche Oberfläche, die in privaten und öffentlichen (für alle Kinder) Bereichen unterteilt ist. [01]

In der ersten Phase stellen die Spielenden ihre Figuren dem Teleporter vor (siehe Abb.12). Durch ihn können ihre Avatare durch die Welt reisen und neues über sie erfahren. Die Kinder helfen ihren Elys die Reise durch Bilder, Videos und Musik zu dokumentieren. Sie können z.B. Grußkarten verschicken, um Informationen für schon vergangenen Reisen darauf zu speichern.

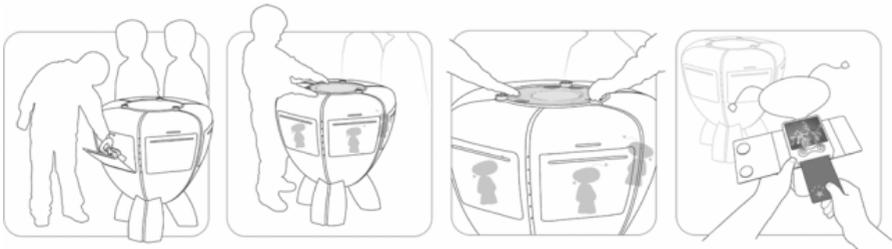


Abb.12 Im ersten Phase des Spiels stellen ihre Elys dem Teleporter vor. [01]

5.2 The Smart Jigsaw Puzzle



Abb. 13 Der Smart Jigsaw Puzzle Assistant benutzt RF-ID Tags, um Puzzleteile ihrer richtigen Position zuzuordnen. [10]

Nach erfolgreicher Umsetzung von klassischen Kartenspielen entwickelte das Team von Jürgen Bohn ein Computer gesteuertes Puzzle-Assistenten [10]. Es handelt sich um eine Tabletop Anwendung (siehe Abb. 13), die RF-ID Tags benutzt, um Puzzletei-

le ihrer richtigen Stelle zuzuordnen. Jedem der Puzzlestücke wird ein kleiner RF-IF Tag eingebettet. Die Anwendung läuft auf einem ganz normalen PC oder Laptop und überwacht die Anordnung des Puzzles im Ganzen. Wenn man einen Teil einer Reihe schon fertiggeordneter Puzzlestücke auf der Spieloberfläche zuordnen möchte, untersucht der RF-ID Leseteil davon, ob er hingehört. Bei Erfolg wird das Bild auf dem Bildschirm mit dem neuen aktualisiert.

5.3 FantasyA

FantasyA ist ein 3D-Computerspiel, das durch SenToy als Benutzerschnittstelle gespielt werden kann. Der Spieler versetzt sich in der Rolle eines Zauberers, der den Führer seines Volks im Land FantasyA finden muss. Das Spiel hat eine Einführungsphase, in der die vier Herde (Luft, Erde, Feuer und Wasser) und die Schlachten vorgestellt werden. Danach wird der Spieler ins FantasyA-Land gelassen, wo er seine Mission erreichen muss. Natürlicherweise begegnet er vielen Feinden, mit denen er sich auseinandersetzen muss. Die Hauptidee der zusätzlichen Benutzerschnittstelle ist, dem Spieler die Macht zu geben, die Emotionen seiner Spielfigur zu beeinflussen, und somit auch die Taten seiner Gegner. SenToy kann die sechs Grundemotionen (Siehe Abb. 14) - Freude, Angst, Traurigkeit, Wut, Schadenfreude und Überraschung, vermitteln. In den Körpergliedern der Puppe sind Sensoren angelegt. [03][04]



Abb. 14 Nach der Position der Gliedern wird durch Sensoren unter sechs Grundemotionen unterschieden (von links nach rechts: Freude, Angst, Traurigkeit, Wut, Schadenfreude, Überraschung). [03]

5.4 KnightMage

Das Rollenspiel KnightMage findet in einem Kerker statt. Unter den Schatzmengen gibt es darin auch eine Unzahl von Monsters und Ungeheuern, die besiegt werden müssen. Die Spieler müssen sich einander gegen die Gefahren helfen, sind jedoch Gegner im Bezug auf die gesammelten Schätze. KnightMage ist auf STARS Plattform umgesetzt worden. D.h. es verfügt über ein sich dynamisch änderndes Riesenspielfeld (der berührungsempfindliche Bildschirm). Auf der Wand wird eine Übersicht der schon entdeckten Spielfläche projiziert. Auf PDAs werden die geheimen Informationen jedes Spielers nur für ihn selbst wie auf Abb. 15 gezeigt: Goldbestand, Kraft, Entdeckungen usw. [05]

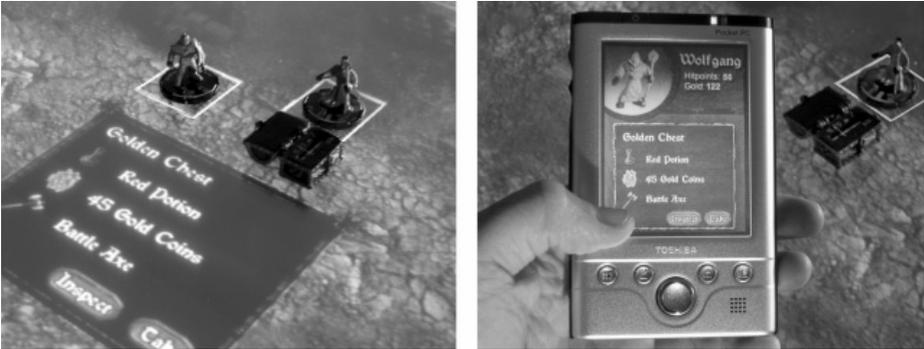


Abb. 15 Die öffentliche (links) und die private (rechts) Sicht einer Spielfigur in KnightMage [05]

5.5 Monopoly®

Die STARS Version von Monopoly bietet im Gegensatz zu dem traditionellen Spiel die Bequemlichkeit der sich rotierenden Spielfläche, sodass sie an dem Blickwinkel des jeweiligen Spielers optimal angepasst ist (siehe Abb. 16). [05]

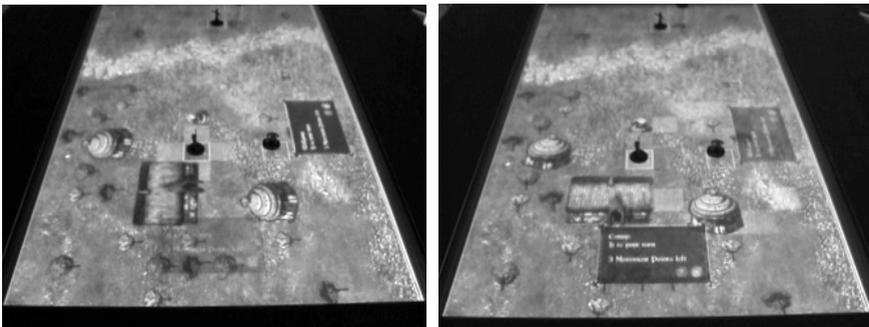


Abb.16 Beispiel für sich dynamisch am Blickwinkel des Spielers anpassende Spieloberfläche der STARS Plattform [05]

Die verschiedenen Aktionen werden gleich am Bildschirm dargestellt. Das Geld ist auch digital präsentiert. Als zusätzliche Funktionalität dazu gibt es auch ein Paar statistische Funktionen. Auf dem Wandbildschirm wird ständig ein Diagramm über den finanziellen Wachstum eines jeden Spielers aktualisiert, was die strategische Planung zum Kaufen oder Verkaufen viel leichter und deswegen amüsanter und angenehmer macht. Die Spieler können sich einander Geld ausleihen, um heimliche Allianzen gegen andere zu entwickeln. Diese Option gebraucht die Funktionalität der PDA-Schnittstellen.

5.6 CYSMN

Can You See Me Now ist ein Spiel für Online- (Players) und mobile Nutzer (Runners) auf den Straßen [08]. Runners und Players haben eine gemeinsame Karte als Spieloberfläche. Die wird online und auf den Handhelds der Runners dargestellt. Im Experiment wurden die Compaq iPAQs benutzt. Sie verfügten über 802.11b WLAN und einen GPS Empfänger. Die Runners, die nicht von den Players gefangen werden dürfen, haben auch eine globale Sicht mit allen Avataren (siehe Abb. 17). Ihre Koordinaten werden durch den GPS Empfänger bestimmt und per WLAN dem Server verschickt. Die Spieler zu Hause müssen die Runners entdecken und fangen, wobei sie nur über eine lokale Sicht verfügen, die nur die nächst gelegenen Avataren zeigen. Ein Läufer zählt genau dann als gefangen, wenn der Avatar eines Spielers weniger von 5 Meter davon entfernt ist. Ein Echtzeitaudiokanal von den Läufern zu den Online Teilnehmer verstärkt den Eindruck für Realität und hilft auch natürlich bei der Bestimmung eines Standortes. Die Läufer tauschen sich durch Walkie-Talkies aus. Die Online Spieler sprechen miteinander durch einen Textkanal.

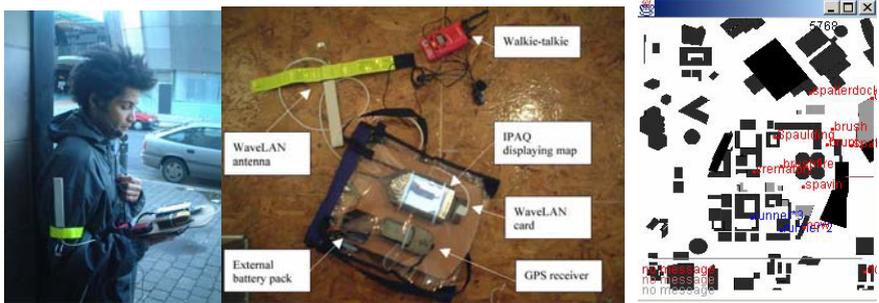


Abb. 17 Ein „Runner“ (links), seine Ausrüstung (Mitte) und die 2D-Darstellung des Stadtplans auf seiner Compaq iPAQ (rechts). [08]

5.6 Coast Puzzle

Coast Puzzle ist ein Tabletopspiel, das mehreren Spielern die Möglichkeit gibt, zusammen ein Puzzle zu ordnen. Sie alle haben eine gemeinsame Sicht, umgesetzt auf das MERL DiamondTouch Plattform, und dürfen gleichzeitig Puzzleteile verschieben, markieren und auf ihren richtigen Platz legen. Somit vereinigt Coast Puzzle die gelassene Atmosphäre eines klassischen Brettspiels mit den Vorteilen einer elektronischen Oberfläche, wo z.B. keine Teile verloren gehen können. [05]

5.7 MemoSet

MemoSet ist eine technische Version vom klassischen Spiel Set, wo 12 Karten auf einem Brett gelegt worden sind. Jede Karte enthält ein gleiches von insgesamt 3 verschiedenen Farben und Formen. Ein Set besteht aus drei Karten, wobei jedes Attribut davon entweder gleich oder verschieden von allen Karten ist. Bei der Computer Ver-

sion wurde die Anzahl der Karten auf 20 erhöht und die Karten werden als zugedeckt initialisiert. Jeder Spieler darf bis zu drei Karten aufmachen, wobei sie dann offen für alle Spieler bleiben. Das Spiel wurde auf dem MERL DiamondTouch Plattform realisiert und erlaubt hiermit bis zu 4 Spielern wie auf Abb. 18. Es ist in Smalltalk implementiert worden und benutzt die COAST groupware Framework. COAST unterstützt die Umsetzung verschiedenster Spiele, da es fertige Datenmodelle zur Verfügung stellt, um die technische Details dem Programmierer zu ersparen. [05]



Abb. 18 Memoset Implementierung für das Multi-Spieler Plattform MERL Diamond-Touch erlaubt bis zu 4 Spieler auf einmal. [05]

6 Fazit und Ausblick

Die Entwicklung neuer Benutzerschnittstellen ermöglicht neue Arten von Interaktion zwischen den Nutzern und dem Rechner, die insbesondere bei Computerspielen von großem Interesse sind. Diese Entwicklung schafft unvorhergesehenen Wirkungen auf die Arbeit des Anwenders. Da der Computer als Vermittler zwischen dem Menschen und seiner Phantasiewelt schon an seinen Grenzen liegt, forscht man nun nach Methoden zur Vermittlung zwischen ihm und des Menschen. Technologien, die auf RFID, Sensoren oder Roomware beruhen, erscheinen sehr passend, da sie die Mobilität und die Abhängigkeit der Spieler von ihrer Umgebung sichern. Die Benutzerschnittstellen sollen auch noch intuitiv und konsistent sein. Besonders bei Spielen ist der Unterhaltungsfaktor am höchstens zu schätzen. Auch insbesondere dann, wenn man sie als Lernmittel einsetzen möchte. Ihre Intuitivität (am Beispiel einer berührungsempfindlichen Oberfläche) erhöht auch die Altersreichweite der Anwender (am Beispiel von Ely The Explorer). Eine sehr wichtige Eigenschaft und Anforderung an Spiele ist eine Multiplayeroption, die den Spielen einen sozialen Kontext hinzufügt. Die neuen Benutzerschnittstellen lassen auch gewisse Einschränkungen in der Kommunikation umgehen (z.B. bei behinderten Menschen). Im Moment handelt es noch um Forschungsprototypen und die Zeit wird zeigen, ob sie das Interesse des breiten Publikums wecken werden.

Literatur

- [01] Africano D., Eriksson S., Lindbergh K., Lundholm P., Nilbrink F., Ely the Explorer: A Multi-User Interactive Play System to Promote Collaborative Learning, In: Proceedings of IDC2003, 7/2003 Preston, UK
- [02] Rekimoto J., Wang H., Sensing GamePad: Electrostatic Potential Sensing for Enhancing Entertainment Oriented Interactions, In: CHI 2004, April 24-29, 2004, Vienna, Austria
- [03] Paiva A., Prada R., Chaves R., Vala M., Bullock A., Andersson G., Höök K., Demo: Playing FantasyA with SenToy, In: ICMI '03, November 5-7, 2003, Vancouver, British Columbia, Canada
- [04] Höök K., Bullock A., Paiva A., Vala M., Chaves R., Prada R., FantasyA and SenToy, In: CHI 2003, April 5-10, 2003, Ft. Lauderdale, Florida, USA
- [05] Magerkurth C., Stenzel R., Prante Th., STARS – A Ubiquitous Plattform for Computer Augmented Tabletop Games, In: Peter Ljungstrand, Jason Brotherton (Ed.): Video Track and Adjunct Proceedings of the Fifth International Conference on Ubiquitous Computing (UBICOMP '03), Seattle, Washington, USA, Oct. 12-15, 2003
- [06] Mazalek A., Davenport Gl., A Tangible Platform for Documenting Experiences and Sharing Multimedia Stories, In: ETP '03, 7/11/2003 Berkley, California, USA
- [07] Magerkurth C., Memisoglu M., Engelke T., Streitz N., Towards the Next Generation of Tabletop Gaming Experiences, In: Proceedings of Graphics Interface 2004, May 17-19, London, Ontario
- [08] Flintham M., Anastasi R., Benford St., Hemmings T., Crabtree A., Greenhalgh Ch., Rodden T., Tandavanitj N., Adams M., Row-Farr J., Where On-Line Meets On-The-Streets: Experiences With Mobile Mixed Reality Games
- [09] Pape S., Dietz L., Tandler P., Single Display Gaming: Examining Collaborative Games for Multi-User Tabletops
- [10] Bohn J., The Smart Jigsaw Puzzle Assistant: Using RFID Technology for Building Augmented Real-World Games
- [11] http://www.smartboard.co.uk/company/mediacenter/press/releases/dvit_release.asp
- [12] <http://www.smarttech.com/DViT/>
- [13] <http://www2.smarttech.com/st/en-US/Products/SMART+Boards/Overlays/>
- [14] <http://www2.smarttech.com/st/en-US/Products/SMART+Boards/Overlays/Software.htm>
- [15] http://www2.smarttech.com/NR/rdonlyres/D4873BCF-50E9-4DB3-9406-BDCAC2782922/0/sbpd_brochure.pdf
- [16] http://www.smartopschool.nl/images/dvit_logo.jpg
- [17] http://www.really-sun.com.cn/newproduction/image/DViT_splash21.jpg
- [18] <http://www.merl.com/projects/images/DTApplications.jpg>
- [19] <http://de.wikipedia.org/wiki/RFID>
- [20] <http://wlan.mail2000.net/wlan.jpg>
- [21] <http://www.handystotal.de/produkte/umts-handy.jpg>
- [22] <http://www.glossar.de/glossar/images/siemens-multimobile.gif>
- [23] http://www.dh.umu.se/fprw/images/projects/2002/L/2002_id_02_L_diana_a.jpg
- [24] http://www.ipsi.fraunhofer.de/ambiente/projekte/projekte/Bilder/interactable_wall_chair.jpg
- [25] <http://www.smartboard.de/products/index.asp>
- [26] <http://www.merl.com/projects/DiamondTouch/>
- [27] <http://www.smarttech.com/company/aboutus/coverage/dvit.asp>
- [28] http://www.wsv.de/fvt/funknavi/gpsvt_1/gpsvt_1.html
- [29] <http://www.informationsarchiv.net/statisch/wlan/>
- [30] http://www.nm.informatik.uni-muenchen.de/Vorlesungen/ws0405/vs_skript/verteilte_systeme.pdf

Berührbare Benutzerschnittstellen und Daten zum Anfassen

Stefan Freund

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`freunds@informatik.uni-muenchen.de`

Zusammenfassung. Wie wird der moderne Arbeitsplatz in der Zukunft aussehen? Werden die Menschen in allen Arbeitsbereichen bald nur noch vor dem Computer sitzen und ihre Arbeit mit Maus und Tastatur verrichten? In vielen Fachgebieten benötigen Benutzer mehr als nur grafische Benutzerschnittstellen, denn der Mensch besitzt auch viel mehr motorische Möglichkeiten als nur zu zeigen, z.B. drücken, drehen, schieben, neigen usw., die er bei den konventionellen grafischen Benutzerschnittstellen nicht einbringen kann.

Die Erfahrung, Präzision und Geschwindigkeit mit einem für seine Arbeit optimierten Fachwerkzeug sind es, die einen Experten so effizient arbeiten lassen. Bei der Umstellung auf eine ausschließlich grafische Benutzerschnittstelle gehen sowohl die Funktionalität des Fachwerkzeugs als auch die Fähigkeiten des Benutzers verloren. Bei der Entwicklung von berührbaren Benutzerschnittstellen (BBS) sieht man davon ab immer mehr Objekte zu virtualisieren. Stattdessen werden physische Objekte, z.B. die Fachwerkzeuge, mit Informationen und Sensoren ausgestattet. Dies ermöglicht dem Benutzer eine intuitivere und unterstützte Bedienung.

Im Rahmen dieser Arbeit werden Entwurf und Entwicklung von berührbaren Benutzerschnittstellen und Daten zum Anfassen vorgestellt. Die Möglichkeiten der Eingabe sowie der Ausgabe werden an Hand existierender Projekte beleuchtet. Abschließend folgt ein Ausblick auf zukünftige Anwendungsgebiete und Anwendungsformen.

1 Einleitung

Berührbare Benutzerschnittstellen (BBS) besitzen im Vergleich zu grafischen Benutzerschnittstellen einen geringeren Bekanntheitsgrad. Die beiden berühmtesten und erfolgreichsten Benutzerschnittstellen sind allerdings berührbar. Das prominenteste Beispiel einer BBS ist die Tastatur, die sich aus der Idee und dem Design der früheren Schreibmaschine entwickelt hat. Nach den Lochkarten war dies die erste wirkliche Benutzerschnittstelle. Ihr Erfolg ist größtenteils damit erklärbar, dass sie das akzeptierte Konzept der Schreibmaschine widerspiegelte und dass zu dieser Zeit noch keine anderen Eingabemöglichkeiten benötigt wurden. Anfangs bestanden die Anwendungen am Computer zum größten Teil aus Text- und Tabellenbearbeitungsprogrammen (z.B. VisiCalc). Als nach und nach grafische Benutzeroberflächen für die Programme

geschaffen wurden (die ersten, entwickelt von Xerox, erschienen 1981), kam der Durchbruch eines schon 1964 von Douglas Engelbart entwickelten Eingabegerätes (siehe Abb. 1). Der "X-Y-Positions-Anzeiger für ein Bildschirmsystem", heute gemeinhin wegen seines Aussehens als "Maus" bezeichnet, erlaubte dem Benutzer erstmals die zweidimensionale Koordination seiner Hand auf den Computer zu übertragen [1, Abschnitt 1].



Abb. 1: Prototyp der Computermaus [2]

Dadurch war es für den Anwender möglich Zeigeoperationen, die mit der Maus ausgeführt wurden, dank der ständigen Abtastung der Sensoren der Maus fortlaufend am Bildschirm zu beobachten. Inspiriert von diesem Erfolg wurden auch andere berührbare Benutzerschnittstellen entwickelt. Diese fanden allerdings keine vergleichbare Akzeptanz unter den Benutzern. Ein Beispiel hierfür sind z.B. "Trackballs" oder "Touchpads" für Laptops, die sich nicht gegen die Maus durchsetzen konnten, da diese schon allgemein als Standard angesehen wurde [3].

Berührbare Benutzerschnittstellen, die sich in einem anderen Bereich schon bewährt hatten und somit von den Benutzern schon gehandhabt werden konnten, wurden für passende Computerprogramme genutzt. Der "Joystick" oder Steuerknüppel, der aus der Luft- und Raumfahrt kam, wurde für Flugsimulatoren und ähnliche Programme genutzt. In ihren Fachbereichen wurden diese Eingabegeräte schnell akzeptiert. Dies führte aber nicht zu einer größeren Verbreitung und Verallgemeinerung dieses Eingabegeräts, sondern zu einer weiteren Spezialisierung.

2 Definition von berührbaren Benutzerschnittstellen

Bevor die Anforderungen und Funktionen einer BBS diskutiert werden können, müssen die einzelnen Bestandteile und die Benutzerschnittstelle als Ganzes definiert werden. Benutzerschnittstellen in diesem Feld wurden von den Forschern mit vielen verschiedenen Namen versehen [4, Abschnitt 1]. Nach und nach einigte man sich auf den Begriff: „tangible user interfaces“ (auf Deutsch: „berührbare Benutzerschnittstellen“, abgekürzt als BBS). Die Basis einer BBS beschreiben Ishii und Ullmer als Objekte, die eine physische Repräsentation von digitalen Daten darstellen [5, Abschnitt 1].

Eine weitere Definition geht einen Schritt tiefer und beinhaltet Benutzerinteraktion. Hier werden BBS von Holmquist, Schmidt und Ullmer als physische Objekte beschrieben, die als veränderbare Darstellung von Informationen dienen [4, Abschnitt 4]. Das Objekt registriert die Veränderung über Sensoren und überträgt sie an den Rechner. Die digitalen Informationen werden entsprechend dem Benutzereingriff interpretiert. Dabei wird dem Benutzer die Illusion gegeben keine Eingabe über die Benutzerschnittstelle gemacht zu haben, sondern die Information direkt verändern zu können (siehe Abb. 2) [5, Abschnitt 6].

Auch die grafischen Benutzeroberflächen erlauben die Manipulation von digitalen Daten. Der entscheidende Unterschied ist allerdings, dass die berührbaren Benutzerschnittstellen mit den digitalen Informationen stets verbunden sind. Bei einer Veränderung eines Attributs des Objekts durch den Benutzer, z.B. bei Veränderung der Position, werden die digitalen Daten ständig angepasst. Veränderungen an den digitalen Informationen führen wiederum zu einer physischen Ausgabe, z.B. selbstständiger Neupositionierung. Der Benutzer muss folglich nicht den Umweg über eine digitale Schnittstelle machen, sondern kann direkt mit seiner Umwelt interagieren, indem die BBS als kombiniertes Ein- und Ausgabegerät benutzt wird [6, Abschnitt 2].

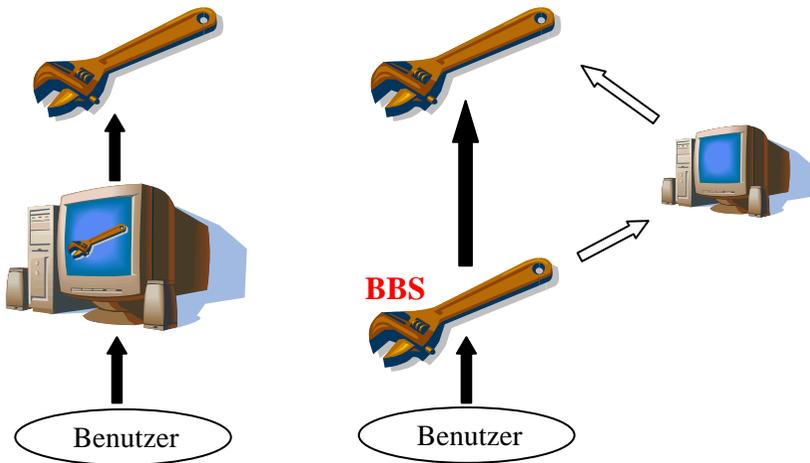


Abb. 2: Wahrnehmung des Benutzers im Vergleich

2.1 Notwendigkeit neuer berührbarer Benutzerschnittstellen

Das Design neuer BBS ist keine triviale Aufgabe, selbst die Maus wurde nach ihrer Veröffentlichung nicht als sinnvoll erachtet. Es gab schlichtweg noch keinen Anwendungsbereich für ein derartiges Eingabegerät, da sich die Programme zu dieser Zeit auf Text- und Tabellenverarbeitung beschränkten. Ein passendes Anwendungsgebiet zu finden und auch Benutzer, die die BBS als sinnvoll und effektiv ansehen und auch benutzen werden, ist eines der Hauptprobleme bei der Entwicklung von BBS. In einigen Fachbereichen wurde dies realisiert, indem schon bestehende fachspezifische Ein- und Ausgabegeräte, die den Benutzern schon vertraut waren, mit dem Computer verbunden wurden. In vielen Bereichen werden grafische Benutzerschnitt-

stellen entwickelt, die zum Teil vorher verwendete Geräte lediglich virtualisieren. Damit wird dem Benutzer ein Anwendungsprogramm in die Hand gegeben, in dem alle bisherigen Eingabemöglichkeiten weiterhin bestehen. Die Eingabe selber erfolgt aber nur noch über Tastatur und Maus. Dies erscheint zunächst nicht als negativ, denn der Benutzer kann schließlich alle Aktionen, die er früher ausgeführt hat, mit Maus und Tastatur durchführen. Doch der Mensch besitzt viel mehr Möglichkeiten, rein mechanisch, als nur zu zeigen. Ein Tontechniker dreht und schiebt Regler an seinem Mischpult (siehe Abb. 3), ein Designer zieht und dehnt eine Modelliermasse und ein Juwelier neigt einen Edelstein und drückt ihn an eine Schleifmaschine.



Abb. 3: Grafische Oberfläche eines virtuellen Tonmischpults [7]

Bei vielen Handlungsabläufen kann die Produktivität und Präzision des Arbeiters auf die Erfahrung und Fähigkeit mit seinem Fachwerkzeug zurückgeführt werden. Bei einer Umstellung auf eine rein grafische Benutzerschnittstelle verliert man somit die Kenntnisse des Benutzers und auch die Funktionalität und Designaspekte des Fachgeräts [5, Abschnitt 3]. Auch gleichzeitige und schnelle Eingaben können in einer solchen grafischen Benutzerschnittstelle kaum nachvollzogen werden, da der Benutzer nur einzelne Aktionen nacheinander ausführen kann und dabei einen Großteil seiner Konzentration auf den Eingabevorgang selbst legen muss. Der Mensch, der über vielfältige und effektive Möglichkeiten verfügt mit seiner Umwelt in Kontakt zu treten, wird somit gezwungen seine normalen Arbeits- und Handlungsabläufe auf eine Vielzahl von Zeigeoperationen umzustellen. Komplexe menschliche Handlungen mit physischen Geräten grafisch zu modellieren ist deshalb keine zufriedenstellende Lösung. Eine erweiterte Benutzerschnittstelle wird benötigt, damit der Benutzer zwar den Nutzen der Optionen des Computers hat aber nicht die Einschränkung und Behinderung in der Ein- und Ausgabe.

Anwendungsfelder für die Entwicklung von BBS sind durchaus gegeben, doch die Umsetzung zu einem Eingabegerät, das dem Benutzer einen wirklichen Mehrwert bei den Arbeitsabläufen schafft, macht den Erfolg einer BBS aus. Auch die Forschung und Entwicklung von neuen, allgemein verwendbaren BBS wächst stark, obwohl Maus und Tastatur immer noch als Standards unangefochten sind. Es eröffnen sich auch neue Bereiche für BBS z.B. 3D Menüs, Betriebssysteme und virtuelle Realität, für die zumindest die Maus einen Nachfolger als Eingabeschnittstelle sucht.

2.2 Entwicklungsansätze von berührbaren Benutzerschnittstelle

Nachdem das Konzept für eine neue BBS gefunden ist, müssen noch viele Designansätze geklärt werden. Zunächst stellt sich die Frage, welche der Eingabe- und Ausgabekanäle man digital oder physikalisch darstellt. Bei einer grafischen Benutzerschnittstelle wird, wie erwähnt, sämtliche Information digital dargestellt, während die Eingabe fast nur über Tastatur und Maus erfolgt. Bei der Entwicklung einer BBS bleibt aber dem Erfinder der BBS die Wahl, wie er Informationen darstellen möchte [8, Abschnitt 2].

Bei der Wahl der Darstellung sollte immer der Standpunkt des Benutzers beachtet werden. Eine eindeutige Verbindung der digitalen Information und der physikalischen Objekte stellt den Grundstein einer erfolgreichen BBS dar [9, Abschnitt 1]. Die Frage, die aber immer im Vordergrund steht, ist: Ist es für den Anwender intuitiv und verständlich, dass er mit dieser Eingabeoperation bestimmte Aktionen beim Rechner auslöst? Dies kann am Beispiel der Maus gut erläutert werden. Bewegt ein Benutzer die Maus nach rechts, so erwartet er, dass sich auch der Mauszeiger am Bildschirm entsprechend nach rechts bewegt. Eingaben zur Veränderung der Positionierung und Ausrichtung können sehr erfolgreich physikalisch dargestellt werden. Hierbei muss allerdings auf die motorischen Fähigkeiten eines Menschen geachtet werden, denn manche Bewegungen können nicht gleichzeitig, schnell oder beliebig lange ausgeführt werden können.

Außerdem gibt es noch Eingaben, die ihrer Art nach physikalisch nicht greifbar sind, wie z.B. Farbe oder Zeit. Eine physikalische Eingabe eines solchen Attributs ist dennoch möglich. Man verwendet Metaphern für die nicht greifbare Information [6, Abschnitt 3]. Bei dieser Eingabe muss sehr darauf geachtet werden, dass der Benutzer diese Aktion auch selbst erkennen, abschätzen und durchführen kann. Man könnte bei der Veränderung von Farbe z.B. eine Malpalette und einen Pinsel anbieten, mit dem der Benutzer das Objekt berühren kann. Der Pinsel stellt hier eine Metapher für das Auftragen von Farbe dar. Falls dies aber eine Aktion ist, die nicht sehr häufig vorkommt, erscheint diese Lösung zu umständlich. In diesem Fall könnte auf eine grafische Benutzerschnittstelle zurückgegriffen werden, in der z.B. ein Dialog zur Farbauswahl geöffnet wird, auf den der Benutzer zugreifen kann. Bei der Verwendung von Metaphern muss allerdings genauestens überlegt werden, wie der Benutzer diese auffasst, denn oft sind Metaphern kulturell oder fachspezifisch verschieden.

Auch bei der Ausgabe muss entschieden werden, welche Informationen grafisch und welche physikalisch übermittelt werden. Bei der Rückgabe von komplexen Operationen, z.B. der Bestätigung einer Speicherung, ist eine physikalische Rückgabe für den Benutzer nicht verständlich. In diesen Fällen wird meistens auf eine visuelle Darstellung des Ergebnisses zurückgegriffen.

Die physikalischen Objekte einer BBS stehen dem Benutzer fortlaufend zur Verfügung. Damit wird der Anwender erwarten, dass auch seine Eingaben und Veränderungen fortlaufend übertragen werden [8, Abschnitt 2]. Eine BBS muss fortlaufend Eingaben empfangen und verarbeiten können, während eine grafische Benutzerschnittstelle dem Benutzer die Interaktionsdialoge auch einfach entziehen kann. Zusätzlich gibt es einen viel höheren Freiheitsgrad bei einer BBS. Beim Design muss der Entwickler der BBS auf jede erdenkliche (bzw. sinnvolle) Interaktionsform eingehen, die mit physikalischen Objekten ausgeführt werden kann. Bei grafischen

Benutzerschnittstellen hingegen sind die Interaktionsmöglichkeiten viel geringer und können bei der Entwicklung genau festgelegt werden.

Es muss weiterhin darauf geachtet werden, dass die physischen Objekte der BBS sich situationsbezogen anders verhalten müssen. Die physischen Objekte werden oft nicht nur durch ihre physikalische Gestalt sondern auch durch ihr Umfeld bestimmt. Somit besitzt ein Objekt seinem Umfeld, also seinem Kontext, entsprechende andere Verhaltensweisen [8, Abschnitt 2].

2.3 Funktionalitätsprobleme und Lösungen

Bei der Implementierung der BBS ist das Hauptproblem die Erfassung der Benutzereingaben und die Verarbeitung dieser Informationen. Es erfolgt eine fortlaufende Kommunikation zwischen Computersystem und den physikalischen Objekten der BBS. Diese können fortlaufend auch von mehreren Benutzern verändert werden.

Ein zusätzliches Hindernis sind die noch fehlenden standardisierten Ein- und Ausgabetechnologien (z.B. bei magnetischen Sensoren oder RFID Chips). So muss der Entwickler die BBS bis ins Detail an die verwendete Technik anpassen [10, Abschnitt 1]. Damit beschränken die zu Verfügung stehenden Ein- und Ausgabetechnologien die Möglichkeiten der BBS stark.

Um dies zu vermeiden wurden mehrere Ansätze verfolgt eine gemeinsame Schnittstelle für die Entwicklung und Umsetzung BBS zu schaffen. Diese so genannten Werkzeugkästen (Toolkits) für BBS geben dem Konstrukteur von BBS die Möglichkeit BBS auf einer höheren Ebene zu entwerfen. Die Funktion der BBS kann, ohne auf die zugrunde liegende Technik eingehen zu müssen, beschrieben werden. Ein weiterer Vorteil dieses Ansatzes ist die Austauschbarkeit der verwendeten Technologien zum Erfassen von Informationen. Die Anforderungen an ein solches Softwarekonzept sind, wie schon erwähnt, die Unterstützung verschiedenster Ein- und Ausgabetechnologien und Möglichkeiten fortlaufende Eingaben zu verarbeiten. Außerdem sollte es die Interaktion zwischen mehreren Benutzern ermöglichen, eine klare und einheitliche Beschreibung der physikalischen Objekte beinhalten und Rückmeldungen an den Entwickler und Benutzer geben können [10, Abschnitt 2].

Die Basis einer solchen Softwarestruktur machte Ullmer in dem er den „Token and Constraints“ (auf Deutsch: „Marken und Beschränkungen“) -Ansatz darstellte. Hier wurden die physikalischen Objekte als Marken, die digitale Information darstellen, gesehen. Beschränkungen sind Regionen, auf denen die Marken platziert und verändert werden können. Nach Ullmer kann man alle BBS als eine Ansammlung von interaktiven Oberflächen, zusammengesetzten Strukturen und diesen Marken und Beschränkungen spezifizieren [11, Abschnitt 1].

Der Ansatz der „Marken und Beschränkungen“ wurde von Shaer, Leland, Calvillo-Gamez und Jacob aufgegriffen und in das „TAC paradigm“ weitergeführt [9]. Sie verwenden die von Ullmer definierten Marken und Beschränkungen, erweitern aber deren Bedeutung. In der „Marken und Beschränkungen“-Definition besteht eine BBS aus physikalischen Objekten. Diese Objekte können Marken, Beschränkungen oder beides sein. Das beruht darauf, dass eine Marke in sich selbst eine Beschränkung darstellt. Physisch gesehen beschränkt die Marke durch ihren eigenen Körper den freien Platz und verändert die Umgebung für andere Marken. Zusätzlich können die

physikalischen Objekte noch eine Variable besitzen, die die digitale Information darstellen soll [9, Abschnitt 2].

Mit diesem Ansatz lassen sich Eingaben an BBS standardisiert darstellen (siehe Tab. 1). Die Eingaben werden als Ereignisse gesehen und können somit fortlaufend behandelt werden. Ereignisse werden mit der Marke, bei der sie auftreten, und mit den vorliegenden Beschränkungen dieser Marke eindeutig beschrieben. Der Entwickler der BBS muss nun nur noch die Rückgabe aus der Marke, ihren Beschränkungen und der evtl. vorliegenden Variablen bestimmen. Mit diesem Ereignismodell können nun Eingaben aufgenommen und verarbeitet werden. Ein weiterer Vorteil ist, dass dieses Modell von keiner bestimmten Ein- und Ausgabetechnologie abhängt und somit Schnittstellen für verschieden Technologien anbieten kann.

TAC	Darstellung	Verhalten			
	Marke	Beschränkung	Variable	Aktion	Rückmeldung
#	Physisches Objekt	Physisches Objekt	Digitale Information	Durchgeführte Aktion des Benutzers	Rückgaben an den Benutzer

Tab. 1: Erfassung von Marken und Beschränkungen

2.4 „TUIster“

Eine recht „einfache“ BBS ist der TUIster, der von der Fakultät für Computerwissenschaften an der Universität des Saarlands entwickelt wurde. Der „TUIster“ ist ein Eingabegerät, das beidhändig bedient wird und dem Benutzer erlaubt sich in hierarchischen Strukturen zu bewegen. Er besteht aus zwei Zylindern, die getrennt von einander bewegt werden können. Auf dem einen Zylinder befinden sich sechs Bildschirmflächen. Drei davon sind immer aktiv und für den Benutzer sichtbar. Mit der Rotation der Seite des TUIsters, die mit den Bildschirmen besetzt ist, bewegt der Benutzer sich in einer Liste von Optionen. Mit Drehen der anderen Seite, also des Griffs, wählt er ein Menü aus bzw. verlässt es wieder. Die Informationen des Benutzers übermittelt der TUIster per Bluetooth [12, Abschnitt 8].

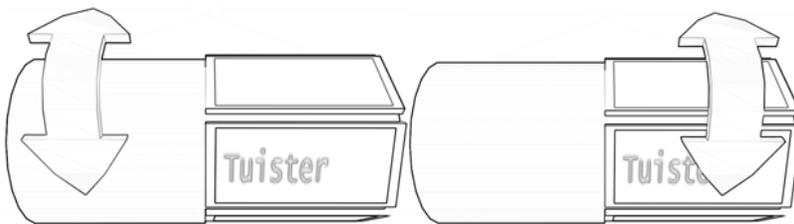


Abb. 4: Eingabemöglichkeiten am „TUIster“ [12]

Das Anwendungsfeld des TUIsters ist vielseitig und bietet die Möglichkeit TUIster mit den verschiedensten Anwendungen zu kombinieren. Er ist portabel und leicht zu handhaben. Benutzer, die oft mit Listen und hierarchischen Menüs umgehen müssen, könnten durchaus einen Mehrwert durch seine Anwendung erzielen.

Der TUIster kann als BBS aufgefasst werden, da er das digitale hierarchische Menü als physisches Objekt widerspiegelt. Bei einer Veränderung an einem der beiden

Zylinder gibt der TUIster diese Information weiter und passt die Bildschirme entsprechend an. Die Rückkoppelung mit dem Benutzer erfolgt allerdings ausschließlich visuell über die Bildschirme auf der rechten Seite (siehe Abb. 4). Der TUIster besitzt zusätzliche Sensoren um die Daten an den Bildschirmen korrekt anzuzeigen, je nachdem wie herum er gehalten wird.

Die Bewegung in der Auswahlliste, durch Drehen des rechten Zylinders, ist für den Anwender intuitiv, da er mit der Drehbewegung nur seiner Leserichtung folgt. Ganz anders ist allerdings die Auswahl eines Menüs. Die Auswahl erfolgt intuitiv über eine Zeige- oder Drückaktion. Hier wurde eine Metapher für den Benutzer eingeführt. Bei der Auswahl eines Menüs zieht der Benutzer dieses mit der Drehung des linken Zylinders fest, wie z.B. bei einer Schraube. Das Verlassen des Menüs entspricht folglich einer Drehung in die andere Richtung und soll eine Lockerung darstellen [13, Abschnitt 5]. Die verschiedenen Eingabeoptionen können mit dem „Marken und Beschränkungen“-Ansatz wie in Tab. 2 beschrieben werden.

TAC	Darstellung		Verhalten		
	Marke	Beschränkung	Variable	Aktion	Rückmeldung
1	rechter Zylinder	keine	Position in der Liste	Nach unten drehen	Nächster Punkt der Liste erscheint auf dem Bildschirm
				Nach oben drehen	Vorheriger Punkt der Liste erscheint auf dem Bildschirm
2	linker Zylinder	keine	keine	Nach unten drehen	Ausgewählte Liste wird verlassen und der Punkt der überliegenden Ebene dargestellt.
				Nach oben drehen	Aktueller Punkt wird ausgewählt, bzw. geöffnet falls es sich um eine weitere Liste handelt.

Tab. 2: Marken und Beschränkungen beim „TUIster“

Die Aufnahme der Eingabe erfolgt über magnetische und optische Sensoren. Diese befinden sich im Inneren des rechten Zylinders. Die magnetischen Sensoren und 3D Beschleunigungssensoren können die absolute Position des TUIster bestimmen, während ein optischer Sensor die Position der beiden Zylinder zueinander bestimmen kann. Diese Daten werden fortlaufend per Bluetooth an einen Rechner geschickt. Dieser bestimmt daraus den dem Benutzer nächsten Bildschirm um dort eine Rückgabe entsprechend der Benutzeraktion auszugeben.

3 Aufnahme der Benutzeraktionen

Im Gegensatz zu einer grafischen Benutzerschnittstelle, in der alle Eingabemöglichkeiten klar definiert sind, ist der so genannte Freiheitsgrad bei BBS viel höher. Der Benutzer kann nahezu jede Aktion, die er mit einem physikalischen Gegenstand ausführen kann, an der BBS anwenden. Die BBS muss nicht für alle diese Eingabemöglichkeiten eine Funktion haben, denn schon auf Grund der Form und der Eigenschaften, von D. Normann als „Affordance“ bezeichnet [14], der BBS wird der Benutzer einige Aktionen ausschließen. Sämtliche logischen und sinnvollen Interaktionen sollten allerdings berücksichtigt werden [9, Abschnitt 1].

Die Handhabung sollte so gestaltet werden, dass auf die verschiedenen Eingabemöglichkeiten schnell und intuitiv zugegriffen werden kann. Die Eingaben müssen fortlaufend übertragen werden um möglichst schnell passende Rückkoppelungen an den Benutzer geben zu können.

3.1 Technologien zur Aufnahme der Benutzeraktionen

Die Anforderungen an Sensoren sind bei BBS besonders hoch und vielseitig. Oft hängen die Möglichkeiten einer BBS maßgeblich von der verwendeten Aufnahmetechnologie ab.

Zu messende Eingaben sind z.B. Position, Ausrichtung, Druck, Geschwindigkeit, Beschleunigung, Wärme und viele mehr. Da die heutzutage benutzten Technologien für diese Eingaben unterschiedlich gut geeignet sind, wird oft eine Kombination von verschiedenen Sensoren benötigt.

Magnetische Sensoren

Diese Sensoren können magnetische Felder u.a. über Induktion in Spulen wahrnehmen (siehe Abb. 5). So wird zum Beispiel zur Messung der Ausrichtung eines Körpers im dreidimensionalen Raum ein Magnetfeld mit niedriger Frequenz erzeugt [15, Abschnitt 1]. Der Empfänger nimmt dieses Magnetfeld wahr und kann über Algorithmen seine Position im Raum daraus ableiten. Diese Technologie eignet sich vor allem zum Erfassen von Entfernungen und Positionen. Ein Nachteil dieser Technologie ist allerdings die Störungsanfälligkeit.



Abb. 5: Magnetischer 3D Sensor [16]

Kameras

Bei der optischen Aufnahme von Informationen werden zunächst Bilder von einer Kamera aufgefangen und an einen Rechner übertragen. Der aufwändige Teil dieser Eingabetechnologie liegt beim Auswerten der Bilder. Zur Erkennung von Bewegungen beispielsweise bestimmt der Computer charakteristische Kanten oder Punkte und berechnet (siehe Abb. 6), wo sich diese im nächsten Bild befinden. Über weitere Rechenverfahren können nun auch Bewegungsrichtung, Geschwindigkeit und Beschleunigung über entsprechende Gradienten bestimmt werden. Bei der visuellen Identifizierung vergleicht der Rechner einzelne Teile des Bildes mit einer Bibliothek ihm bekannter Muster. Ist der prozentuale Grad der Übereinstimmung hoch genug, so hat der Computer das Objekt erfolgreich identifiziert. Das entspricht oft einem Lernprozess. Der Computer speichert Informationen über erkannte Objekte und erweitert so seinen Erkennungsgrad [17, Abschnitt 4]. Die visuelle Erkennung eignet sich sehr gut für die Aufnahme von Farben, Formen und Bewegungen. Der Vorteil dieser Technologie ist, dass nur ein Empfänger, die Kamera, benötigt wird. An dem gemessenen Objekt müssen keine weiteren Hilfsmittel vorhanden sein. Dies ist z.B. bei der Erkennung von Personen ein wichtiger Aspekt.

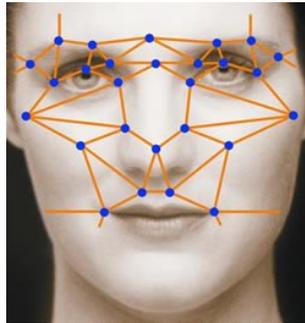


Abb. 6: Punkte zur visuellen Erkennung [18]

RFID

Radio Frequency Identification (auf Deutsch: radio-frequenzbasierte Identifikation) setzt sich aus zwei Bestandteilen zusammen. Zunächst wird eine Antenne benötigt, die Funksignale aussendet. Das Gegenstück dazu ist der Empfänger, i.d.R. ein so genannter Chip (siehe Abb. 7). Er wird durch eine kleine Spule dargestellt, die die Radiowellen aufnimmt und durch Induktion daraus Strom gewinnt um ein Signal zurückzusenden. Dieses analoge Funksignal wird von der ursprünglichen Antenne wieder aufgenommen und mit Hilfe eines Transponders digitalisiert [19, Abschnitt 1]. Bei den RFID Chips unterscheidet man zwischen aktiven und passiven Chips. Passive Chips senden, nach ihrer Aktivierung vom Sender, nur eine weltweit einmalige Identifikationsnummer. Aktive Chips verfügen über eine Stromquelle und einen eigenen Speicher und können somit auf die gesendeten Signale der Antenne reagieren. Der Vorteil der Funk-Erkennung ist, dass keine visuelle Verbindung zwischen Sender und Empfänger bestehen muss. Die RFID Chips eignen sich besonders gut zur Bestimmung der Anwesenheit von Objekten, nicht aber deren Lage oder Orientierung [20, Abschnitt 7].

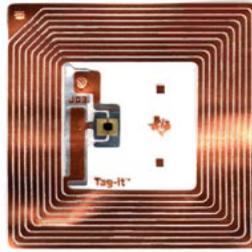


Abb. 7: Passiver RFID Chip [21]

3.2 “DataTiles”

Im Projekt “DataTiles” stellen J. Rekimoto, B. Ullmer und H. Oba eine neue BBS vor, die eine Verbindung von grafischen und berührbaren Schnittstellen darstellt (Siehe Abb. 8). Die Idee war es, digitale Geräte flexibler und einfacher zu gestalten und so dem Benutzer einen intuitiveren Umgang mit Medien zu ermöglichen [22, Abschnitt 1]. Dabei sollten die Vorteile der grafischen Benutzerschnittstellen beibehalten, dem Benutzer aber durch physikalische Module die Handhabung vereinfacht werden.



Abb. 8: „DataTiles“ [23]

Das „DataTiles“ System besteht zunächst aus einem Flachbildmonitor, unter dem ein magnetischer Empfänger für einen digitalen Stift montiert ist (siehe Abb. 9). Auf dem Flachbildschirm befinden sich in regelmäßigen Abständen RFID Sender.

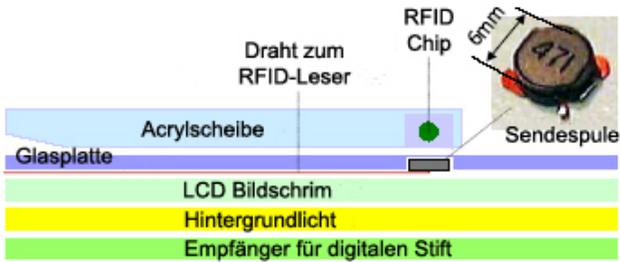


Abb. 9: Querschnitt der „DataTiles“ Arbeitsfläche [22]

Auf dieser Arbeitsfläche können nun einzelne „DataTiles“ abgelegt werden. Diese Acrylscheiben sind durchsichtig und besitzen eine Fußzeile mit einer Beschreibung ihrer Funktion und einem internen passiven RFID Empfänger.

Durch die RFID Chips besitzt jedes „DataTile“ eine einmalige Identität und der Computer kann jedem dieser Medien Daten zuordnen. Wird ein „DataTile“ auf die Arbeitsfläche gelegt, so erkennen die RFID-Sender das Medium und leiten diese Information an den Rechner weiter. Dieser kann aus der Identität des Mediums und aus der Information, welcher RFID-Sender ihm diese Nachricht geliefert hat, eine passende Rückgabe an den Flachbildschirm senden. Auf dem Flachbildschirm werden dann genau unter dem Medium die Daten dargestellt, die es enthält. Der Benutzer bekommt somit die Illusion, die Daten befänden sich direkt auf dem „DataTile“.

Die einzelnen Medien dienen aber nicht nur zur Repräsentation von Daten sondern können auch miteinander interagieren. Dazu gibt es bei „DataTiles“ verschiedene Kacheltypen. Die „Funktionskacheln“ besitzen immer eine schon vordefinierte Funktion, die beim Auflegen aktiviert wird. Beispielanwendungen sind ein Fotoalbum, das einzelne Bilder darstellt, oder ein Videospiele, der einen Film abspielt. Damit der Benutzer diese Aktionen manuell steuern kann, wurden „Parameterkacheln“ eingeführt (siehe Abb. 10). Über diese kann man benachbarte Kacheln steuern und verhindert damit, dass die „Funktionskacheln“ mit zu vielen Eingabemöglichkeiten überladen werden.



Abb. 10: Funktionskachel mit zwei anliegenden Parameterkacheln [23]

Der Benutzer kann mit einem digitalen Stift Eingabefelder der einzelnen Kacheln aktivieren. Auch die Interaktion zwischen mehreren Kacheln wird durch den Stift gesteuert (siehe Abb. 11). Liegt eine „Funktionskachel“ neben einer „Containerka-

chel“, die einen Medienordner darstellt, so können Daten mit einer Stiftbewegung von der „Funktionskachel“ zur „Containerkachel“ geschoben werden. Intern registriert der Empfänger für den digitalen Stift die Bewegung und leitet sie an den Rechner, der daraus die Positionen der beiden Kacheln bestimmen kann. Über die Identitäten der RFID Chips kann er nun die Daten, die der „Funktionskachel“ zugeordnet waren, auch der „Containerkachel“ zuweisen. Der Benutzer bekommt als Rückmeldung nur eine angepasste Darstellung der „Containerkachel“ und ist somit im Glauben die Daten direkt bewegt zu haben.



Abb. 11: Benutzer interagiert mit der Parameterkachel [22]

Durch verschiedene Anordnung von Kacheln können so komplexe Aktionen konstruiert werden. Für den Benutzer sind die Interaktionsmöglichkeiten aber leicht zu verstehen, da er die Daten und Funktionen direkt mit den physischen Kacheln verbindet. Somit kann der Anwender die ihm bekannte Struktur der realen Welt auf die verschiedenen Kacheln übertragen [22, Abschnitt 4].

Die Ausgabe erfolgt bei dieser BBS, wie auch bei vielen anderen Modellen, nur über visuelle Perzeption. Werden Eigenschaften einer „Parameterkachel“ verändert, so verändern sich auch die Bilder der „Parameterkachel“ und der mit ihr verbundenen „Funktionskachel“.

4 Spürbare Ausgabe von BBS

Wie bei den „DataTiles“ beobachtet, ist die visuelle Ausgabe die am meisten verwendete Ausgabemodalität. Dies liegt maßgeblich daran, dass Sehen der primäre Wahrnehmungssinn des Menschen ist und der Benutzer schon gewohnt ist Informationen visuell aufzunehmen. Die Umsetzung der visuellen Darstellung ist auch schon ausgiebig erforscht und wird technisch umfassend unterstützt. Die klassische Darstellung erfolgt an Bildschirmen, wobei auch einfache Flächen über Projektoren zu Bildschirmen verwandelt werden können.

Der Mensch besitzt aber mehr Sinne um Informationen wahrzunehmen, z.B. fühlen oder hören. Damit der Benutzer einer BBS auch wirklich die Illusion hat, dass er die digitalen Daten direkt manipulieren kann, müssen auch die Ausgabekanäle mit denen der realen Welt übereinstimmen. Der Benutzer sieht die BBS als direkte Darstellung der digitalen Informationen und erwartet damit eine exakte Übereinstimmung bei allen Interaktionen [9, Abschnitt 1].

Ein Beispiel für diese Ausgabe ist die „Actuated Workbench“ (auf Deutsch: der gesteuerte Arbeitstisch). Der gesteuerte Arbeitstisch stellt die Verteilung von Objekten im zweidimensionalen Raum dar. Digitale Objekte werden auf der Oberfläche durch Magnete dargestellt. Unter der Arbeitsfläche befinden sich Elektromagnete und magnetische Sensoren. Bei einer Rückgabe des Rechners können so die Magnete entsprechend bewegt werden (siehe Abb. 12) [24, Abschnitt 1]. Würde diese Information nur am Bildschirm dargestellt werden, würden die Positionen der Magnete auf der Arbeitsfläche nicht mehr mit der Position der digitalen Objekte übereinstimmen.



Abb. 12: Aufbau des „Actuated Workbench“ [24]

Der Vorteil von realistischen Rückgaben liegt also auf der Hand. Das Problem ist die technische Umsetzung. Über kleine Motoren lässt sich z.B. eine spürbare Rückgabe an den Benutzer weiterleiten. Diese Technologie wird als „Force Feedback“ (auf Deutsch: Kraft-Rückkoppelung) bezeichnet und ist bei einigen BBS, z.B. Maus und Steuerknüppel, schon durchaus verbreitet. Bei der Ausrichtung und Positionierung im realen Raum sind noch immer große Probleme gegeben, da nicht jedes physische Objekt leicht aus sich heraus bewegt werden kann.

Die auditive Rückmeldung ist im Gegensatz dazu schon weit entwickelt. Diese Technologie wird dennoch eher spärlich eingesetzt, da der Mensch größere Mengen an Informationen auditiv nicht so gut wahrnehmen kann. Ein verbreitetes Anwendungsfeld sind aber ambiente Geräusche, wie z.B. bei einer Fehlermeldung, die dem Benutzer eine Rückmeldung praktisch aus dem Hintergrund geben [5, Abschnitt 6].

5 Daten zum Anfassen

Grundlegend kann zwischen den BBS wie dem „TUIster“ und den „DataTiles“ bzw. „Actuated Workbench“ wie folgt unterschieden werden. Der „TUIster“ erlaubt hier die Manipulation von digitalen Daten, stellt aber keine genaue Repräsentation dieser dar. Die „DataTiles“ hingegen spiegeln die einzelnen Daten wieder und sind somit für den Benutzer anfassbare Daten.

Entscheidet man sich beim Entwurf einer BBS dafür, die digitalen Daten mit physischen Objekten darzustellen, müssen einige Aspekte in Betracht gezogen werden. Zunächst wäre da die physische Gestalt der Daten. Die Form sollte dem Benutzer bereits einen Aufschluss über die Art der Daten geben. Eine Marke, die für ein Video

steht, könnte die Form einer Filmrolle besitzen. Marken mit verschiedenen Funktionen oder Eigenschaften sollte für den Benutzer erkennbar oder wenigstens unterscheidbar sein. Außerdem sollten sie von ihrer Form gut und intuitiv handhabbar sein.

Weiterhin müssen die physischen Beschränkungen, wie beim „Marken und Beschränkungen“-Ansatz, bestimmt werden. Die Umgebung gibt genauso wie die Form, dem Benutzer schon erste Hinweise auf durchführbare Aktionen. Deshalb sollte klar erkennbar sein, welche Bestandteile der BBS dafür gedacht sind mit den Daten zu interagieren und welche nicht. Durch eine sinnvoll gestaltete Umgebung können so schon viele nicht vorgesehene Handlungen unterbunden werden [8, Abschnitt 2].

5.1 „Marble Answering Machine“

Eine der frühesten Ideen zur Verwendung von physischen Objekten zur Repräsentation digitaler Daten hatte Durrell Bishop. Er entwarf die „Marble Answering Machine“ (auf Deutsch: Murmel-Anrufbeantworter) bei der Telefonanrufe durch Murmeln dargestellt werden (siehe Abb. 13) [5, Abschnitt 13].

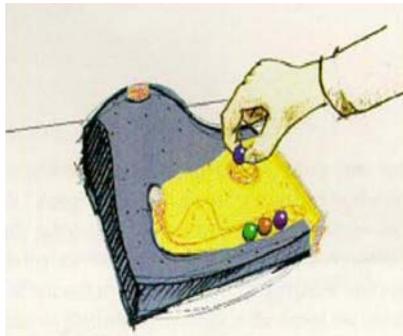


Abb. 13: Design der „Marble Answering Machine“ [5]

Der Anrufbeantworter selbst besitzt zwei Mulden, in die jeweils eine Murmel passt, und eine lange Kerbe, in der Murmeln wie in einer Warteschlange aufgereiht werden können. Geht ein Anruf ein, so wird dieser intern mit der Identität der Murmel verknüpft. Dank der Warteschlange weiß der Benutzer zusätzlich, in welcher Reihenfolge die Anrufe eintrafen. Zum Abspielen der Nachricht muss der Benutzer nur die Murmel in die dafür vorgesehene Mulde legen. Um den Anruf zu erwidern legt er die Murmel in der anderen Mulde ab [5, Abschnitt 13].

Dank des einfachen Designs wäre die Handhabung des Anrufbeantworters für den Anwender intuitiv zu verstehen und deshalb leicht zu erlernen. Die „Marble Answering Machine“ dient oft auch als Beispiel dafür, wie praktisch eine anfassbare Darstellung von digitalen Daten wäre. So könnte z.B. die Murmel und der damit verbundene Anruf einfach eingesteckt und mitgenommen werden.

5.2 „metaDESK“

Die BBS „metaDESK“ wurde am MIT MediaLab entworfen um zu zeigen, wie digitale Information mit physischen Objekten verbunden werden kann. Der „metaDESK“ stellt ein Gelände dar, das mit physischen Modellen von Gebäuden bestückt ist (siehe Abb. 14). Das Computersystem ist sich über den Standort jedes dieser Objekte bewusst. Über eine „aktive Linse“, einen LCD Bildschirm, kann ein virtuelles Bild des Geländes dargestellt werden, in dem die selbst eingefügten Gebäude integriert sind. Eine zweite „passive Linse“ erlaubt dem Benutzer wie mit einer Lupe das Gelände abzufahren und stellt das unterliegende Gelände als dreidimensionales Modell dar.



Abb. 14: Aktive Linse bei „metaDESK“ [25]

Über Kameras und Sensoren erhält das Computersystem Informationen über die Positionen der Gebäude und den Beobachtungswinkel der „aktiven Linse“. Aus diesen Daten kann dann ein virtueller Blick auf das Gelände berechnet werden.

Weiterhin wurde versucht die einzelnen Grundbausteine einer grafischen Benutzerschnittstelle in die BBS zu übertragen. Hierbei wurden neben den Objekten auch Funktionen physisch dargestellt. So wurde z.B. ein Größenregler, in Form eines Schiebereglers, als Konstrukt aus zwei zueinander verschiebbaren Zylindern dargestellt. Im Gegensatz zum grafischen Schieberegler besitzt der physische den Vorteil, dass er an neuen Stellen angelegt werden kann und dort neue Funktionen erfüllen kann [25, Abschnitt 1].

6 Diskussion

BBS sind in vielerlei Hinsicht grafischen Benutzerschnittstellen überlegen. Der Benutzer kann eine intuitivere und umfangreichere Eingabe vornehmen. Dabei wird das Computersystem mehr und mehr unsichtbar für den Benutzer und er bekommt die Illusion die digitalen Daten direkt manipulieren zu können. Außerdem werden die schon erlernten motorischen Fähigkeiten und Erfahrungen des Benutzers genutzt [5, Abschnitt 3].

Beim Entwurf einer BBS muss geklärt werden, welche Informationen physisch oder nur grafisch dargestellt werden. Die Benutzbarkeit einer BBS hängt auch zum großen Teil von einer sinnvollen Verbindung zwischen digitalen Daten und physischen Objekten ab [8, Abschnitt 2]. Je näher die physischen Objekte und die logischen Verbindungen der realen Welt gleichen, desto leichter kommen Benutzer mit der Eingabe zurecht. Die Konstruktion von BBS ist auch auf technischer Ebene keine triviale Aufgabe, da es noch keine standardisierten Ein- und Ausgabegeräte gibt.

Auch wenn die grafische Benutzeroberfläche in der näheren Zukunft noch der Standard der Ein- und Ausgabe sein wird, gibt es heutzutage schon viele Fachbereiche, in denen sich BBS entwickelt haben. Die grafischen Benutzerschnittstellen in diesen Bereichen, wie z.B. der Medizin oder der Landschaftsplanung, waren zu komplex zu bedienen und benötigten sehr gute Computergrundkenntnisse. Die BBS hingegen eröffnen dem Benutzer eine fast reale und deswegen leicht zu handhabende Benutzerschnittstelle. „Illuminating Clay“ (auf Deutsch: Leuchtender Ton) präsentiert geografische Daten auf einem realen Landschaftsmodell aus Ton. Der Benutzer kann die Landschaft manuell verändern, was gleichzeitig vom Computersystem erfasst wird (siehe Abb. 15) [8, Abschnitt 3].



Abb. 15: „Illuminating Clay“ [26]

Literatur

- [1] Wikipedia, Maus (EDV), http://de.wikipedia.org/wiki/Maus_%28EDV%29, 2004
- [2] Prof. Dr. Walter F. Tichy, Informatik I, <http://monet.unibas.ch/~guggisbe/kurs/lek1/bis1970.htm>, 2004
- [3] M. Esser, PC Mäuse, Hamburger Abendblatt, 5.6.1998
- [4] L. Holmquist, A. Schmidt, B. Ullmer, Tangible interfaces in perspective, Pers. Uniquit Comput. 8: 291-293, 2004
- [5] H. Ishii, B. Ullmer, Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, CHI'97
- [6] K. Fishkin, A taxonomy for and analysis of tangible interfaces, Pers. Uniquit Comput. 8: 347-358, 2004

- [7] A. Butz, A. Schmidt, Input and Output – Devices, Taxonomy, http://www.medien.informatik.uni-muenchen.de/fileadmin/mimuc/mmi_ws0405/vorlesung/2004_11_02_f.pdf, 2004
- [8] E. Sharlin, B. Watson, Y. Kitamura, F. Kishino, Y. Itoh, On tangible user interfaces, humans and spatiality, *Pers Ubiquit Comput.* 8: 338-346, 11.12.2004
- [9] O. Shaer, N. Leland, E. Calvillo-Gamez, R. Jacob, The TAC paradigm: specifying tangible user interfaces, *Pers. Ubiquit Comput.* 8: 359-369, 2004
- [10] Scott R. Klemmer, Jack Li, James Lin, James A. Landay, Papier-Maché: Toolkit Support for Tangible Input, CHI 2004
- [11] B. Ullmer, Tangible interfaces for manipulating aggregates of digital information, MIT 2002
- [12] A. Butz, M. Groß, A. Krüger, TUISTER: a multi purpose tangible UI device, <http://www.tuister.com/>, 2004
- [13] A. Butz, M. Groß, A. Krüger, TUISTER: a Tangible UI for Hierarchical Structures, IUI'04
- [14] D. Normann, Design of Everyday Things, New York, 1988
- [15] S. Zachow, J-H. Israel, K. Köchy, Sensor und Kommandotechnologien, <http://cg.cs.tu-berlin.de/~kai/vrmed/node6.html>, 2004
- [16] SLS Pulsed Magnets Group, Magnetic Field Mapping, http://slsbd.psi.ch/timing/Pulsed_Magnets/K1to5/mmap.html, 2004
- [17] C. Schnörr, D. Cremers, T. Kohlberger, Visuelle Objekterkennung – Computer lernen sehen, http://www.uni-mannheim.de/users/presse/forum/p/1_9.html, 2004
- [18] Security.at.work, Pattern Recognition Case Study, <http://www.securityatwork.org.uk/Main/patternCS.htm>, 2004
- [19] Wikipedia, Radio Frequency Identifikation, http://de.wikipedia.org/wiki/Radio_Frequency_Identification, 2004
- [20] AIM Global, What is Radio Frequency Identification? (RFID), http://www.aimglobal.org/technologies/rfid/what_is_rfid.asp, 2004
- [21] W. Hascher, Identifikation mit Mini-Chips, http://www.elektroniknet.de/topics/kommunikation/fachthemen/2003/0021/index_c.htm, 2004
- [22] J. Rekimoto, B. Ullmer, H. Oba, DataTiles: A Modular Platform for Mixed Physical and Graphical Interactions, SIGCHI'01
- [23] J. Rekimoto, DataTiles: A Modular Platform for Mixed Physical and Graphical Interactions, <http://www.csl.sony.co.jp/person/rekimoto/datatile/>, 2004
- [24] G. Pangaro, D. Maynes-Aminzade, H. Ishii, The Actuated Workbench: Computer-Controlled Actuation in Tabletop Tangible Interfaces, UIST 2002
- [25] B. Ullmer, H. Ishii, The metaDESK: Models and Prototypes for Tangible User Interfaces, UIST'97
- [26] C. Ratti, B. Piper, Tangible Infoscapas, <http://web.media.mit.edu/~ratti/tangibleledesign/tangibleledesign21.htm>, 2004

Greifbare Programmierung

Friederike Otto

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
ottof@informatik.uni-muenchen.de

Zusammenfassung Auch wenn wir täglich von immer mehr digitaler Technik umgeben sind, bleibt es doch nur wenigen Menschen vorbehalten selbst Programme zu schreiben. Würde der Akt des Programmierens aus der abstrakten Ebene der Desktop Implementierung entfernt, und in die reale Welt integriert werden, so könnte Programmierung für jeden möglich werden. Gerade Kindern bieten solche Systeme neue Lernmöglichkeiten um ihr logisches Denken zu fördern und spielend mit den Möglichkeiten der digitalen Technik vertraut zu werden. Die Anfänge der Greifbaren Programmierung gehen zurück auf die 1960er und bis heute wird an entsprechenden Systemen gearbeitet. Diese Arbeit gibt einen Überblick über die Ansätze die bisher entwickelt wurden und Systeme die tatsächlich den Weg in die Schulen gefunden haben.

1 Greifbare Programme, warum und für wen? Ein Einblick in die Logo – Philosophie.

Programme greifbar machen, im Sinne von physisch greifbar, aber auch, und das versucht dieser Forschungszeitweig zu erreichen, gedanklich greifbar, also den Prozess zur Manipulation computergesteuerter Geräte zu entmystifizieren und somit der breiten Öffentlichkeit zugänglich zu machen. So soll es zum Ziel werden den Akt des Programmierens aus dem abstrakten Umfeld von Monitor und Tastatur herauszunehmen und in das direkte Umfeld des Nutzers zu transferieren.

Den Grundstein für diese Entwicklung legte Seymour Papert [23] bereits in den 1960`er Jahren mit Gründung der Logo Philosophie [19].

In seiner Theorie des Konstruktivismus stellt er den ungleich höheren Lernerfolg dar, der sich durch „Learning by making“ einstellt. Hier sollen Kinder selbstständig ihre eigenen Ideen umsetzen können, durch den gesamten Prozess der Edit- Compile- Debug Schleife ihr Wissen festigen und jeden Fehler als Möglichkeit erkennen Zusammenhänge besser zu verstehen, denn:

“Life is not about knowing the right answer -(...)- it is about getting things to work!”
(Seymour Papert) [19]

Kinder als Zielgruppe sind nahe liegend, lernen sie schon im Kleinkindalter Wissen zu prüfen und zu festigen, indem sie Situationen aus dem Alltag nachstellen (Puppenspielen, Playmobil, Legobausteine etc.). So ist der Schritt nicht mehr weit Computer in das reale Umfeld der Kinder zu integrieren um Zusammenhänge greifbar zu machen und somit völlig neue Lernmöglichkeiten zu schaffen [20]. Doch diese Möglichkeiten verbaut man ihnen, wenn sich Computerinteraktion auf das abstrakte Level der Desktop Programmierung begrenzt.

Solche Grenzen zu sprengen ist Ziel des Forschungszweiges der Greifbaren Programmierung. Diese Arbeit wird einige Entwicklungen aus diesem Bereich darstellen und durch Beispiele belegen, dass Kinder durchaus in der Lage sind kleine Programme zu implementieren, wenn man ihnen das richtige Werkzeug zur Hand gibt. So sollen die Kinder nicht nur den Dingen das Denken beibringen, sondern durch sie auch selbst zum Denken angeregt werden und neue Denkweisen entwickeln.

2 Bisherige Entwicklungen im Bereich der Greifbaren Programmierung

2.1 Floor Turtle – Geometrie zum Anfassen

Im Zuge seiner Konstruktivismus- Theorie versuchte Seymour Papert (Abb.1) auch die Lernmöglichkeiten an Schulen zu verbessern. So wollte er den Kindern kein abstraktes, und damit schwer zugängliches geometrisches Wissen lehren, sondern sie Geometrie erfahren lassen. Denn wie in allen Lebensbereichen wird auch die Mathematik verständlicher, wenn man sie greifbar erleben kann und so war die Turtle Geometrie geboren.



Abb.1 [6]

Die erste Floor Turtle[1] wurde Anfang der 70er am M.I.T. entwickelt, ein mechanischer Roboter, der via Kabel mit einem Computer verbunden war. Der Roboter, der bewusst auf einen einfachen Befehlssatz (Vor, Zurück, Links, Rechts) beschränkt war, konnte die am Computer eingegebenen Befehle (in der Programmiersprache Logo) verarbeiten und direkt ausführen.

Durch die direkte Reaktion der Floor Turtle (Abb. 2,3) in ihrer realen Welt und der damit verbundenen Identifikationsmöglichkeit der Kinder konnten diese schnell einfache Algorithmen zum Aufbau geometrischer Formen entwickeln. Indem sie selbst „Schildkröte spielten“ wurde es leicht nachzuvollziehen, dass sich beispielsweise ein Quadrat aus dem einfachen Algorithmus – Wiederhole viermal: gehe nach vorn und drehe dich um 90 Grad – zusammensetzt, und schon war ihr erstes Logo Programm geschrieben: [3]

```
TO SQUARE:
```

```
  REPEAT 4 [FORWARD: 50  
            RIGHT 90]
```

```
END
```

So ließ sich der Roboter mittels der prozeduralen Programmiersprache Logo steuern, das heißt, dass die einzelnen Befehlssätze sequentiell von der virtuellen Maschine in der Floor Turtle abgearbeitet wurden.

Doch war es den Erfindern nicht wichtig den Kindern Programmierkenntnisse zu vermitteln, vielmehr war es nun möglich abstrakte Mathematik und Physik spielend in die Umwelt der Kinder zu integrieren („Toys to think with“).

Die Entwicklung der Floor Turtle:



Abb.2 [25]



Abb.3 [24]

Mit dem Siegeszug des PCs in den 80er Jahren fand auch die Floor Turtle ihren Weg in die Schulen. Obwohl bis heute vereinzelt eingesetzt, konnte sie sich auf Grund des hohen Preises, ihrer Unzuverlässigkeit und der mangelnden Möglichkeit komplexere geometrische Formen zu erschaffen, nicht weiter durchsetzen.

2.2 Screen Turtle – Ein Schritt zurück ist ein Schritt nach vorn **Die Turtle auf dem Monitor**

In den 80er Jahren rückte das M.I.T. von ihrer Idee der Greifbaren Programme ein Stück weit ab, und entschied sich auf Grund zahlreicher Vorteile dafür die Turtle auf den Bildschirm zu bringen [2].

Obwohl es sich jetzt nicht mehr um einen fassbaren Gegenstand handelt, soll an dieser Stelle doch kurz auf dieses Kapitel in der Entwicklungsgeschichte Greifbarer Programme eingegangen werden.

Ein Blick auf die Vorteile der Screen Turtle erklärt den ungewöhnlichen Schritt zurück auf den Monitor. So wurde der ursprüngliche Befehlssatz der Floor Turtle (ca.20) übernommen und auf über 100 Instruktionen erweitert [4], womit das implementieren komplexerer Formen möglich wurde (Abb.4). Die Analogie der Schildkröte wurde ebenfalls beibehalten und so werden Befehle immer in Blickrichtung der Turtle ausgeführt, wie der virtuelle Stift auf dem Monitor weiterhin genannt wurde.

Neben der Entwicklungsumgebung veränderte sich auch die zugehörige Programmiersprache Logo, sie wurde toleranter gegenüber den Namen der Primitiven (Forward = For = FD = Ahead), sowie in ihrer Syntax. Die Primitive, denen Parameter mitgegeben werden mussten (z.B. Forward 100) erhielten als Einheit Turtle Steps, die je nach Auflösung des Bildschirms variierten und denen meist die kleinste, am Monitor darstellbare Distanz zu Grunde lag. Des Weiteren gehörten das Einführen von

```

to square
repeat 4 [fd 200 rt 90]
end

to flower
repeat 10 [square rt 36]
end

```

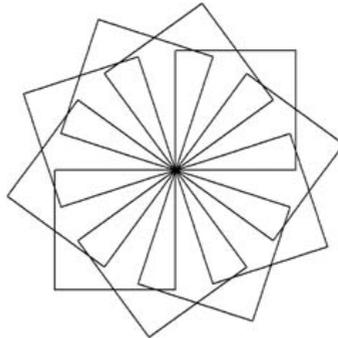


Abb.4 [26]

Vektoren (relative Abstände), sowie verschiedene Farben und Füllmöglichkeiten zu den Neuerungen.

Ganz offensichtlich war die Screen Turtle nutzerfreundlicher und praktikabler als ihr Vorgänger, doch der ursprüngliche Gedanke der Greifbaren Programmierung war verloren gegangen, als sie aus der Alltagsumgebung der Kinder herausgenommen wurde.

Nicht nur die Identifikationsmöglichkeit, also das Nachspielen der Turtle war verloren gegangen, auch wurde die Screen Turtle wieder ein abstraktes Objekt für das nicht klar war, was zum Beispiel links oder rechts bedeutet (aus Sicht der Kinder, oder aus Sicht der Turtle), da sich die Interaktion auf einen zweidimensionalen, Bildschirm beschränkte.

Doch die Erfahrungen die aus diesem Projekt hervorgingen und der hohe Lernerfolg der bei den Kindern zu beobachten war, mobilisierte weitere Forschungsgruppen die Vorteile beider Turtles zu vereinigen.

2.4 Programmierbare Bausteine - Greifbare Programme erobern das Kinderzimmer

Bereits seit 1987 wurde am M.I.T. Media Laboratory an verschiedensten Entwürfen eines autonomen programmierbaren Bausteins gearbeitet

Bei ihren Entwicklungen, wie im gesamten Bereich der Greifbaren Programmierung sahen sie Kinder als Designer und Entwickler eigener Programme und wollten somit ihre Rolle nicht nur auf das manipulieren vorgegebener Systeme beschränken. Doch mit ihrer Idee der programmierbaren Bausteine (P- Bricks) wollten sie noch ein Stück weitergehen als es die bisherigen Entwicklungen erlaubten und setzten sich folgende Entwurfsziele [5]:

- Variable Einsatzmöglichkeiten:

Die Bausteine sollten eine Vielzahl von Einsatzmöglichkeiten abdecken, beispielsweise das Steuern von Robotern, das Gestalten interaktiver Räume oder auch das Durchführen physikalischer Experimente.

- Verschiedenste I/O Modalitäten:

Der P-Brick sollte unterschiedlichste Einflüsse aus der realen Welt verarbeiten können, z.B. durch Berührungs-, Licht- und Temperatursensoren, sowie Infrarotempfänger. Auch die Interaktionsmöglichkeiten mit der Außenwelt wurden durch erweiterte Ausgabemöglichkeiten wie Motoren, Licht, Sound und Infrarot verbessert. So boten sich aus den Kombinationen der Ein- und Ausgabe verschiedenste Einsatzmöglichkeiten.

- Parallelität:

Gerade bei Kindern ist es wichtig die Möglichkeit bereitzustellen, mehrere Dinge gleichzeitig zu kontrollieren, sonst verlieren sie schnell die Lust an ihrem Spielzeug. Daher sollte der P-Brick mehrere Ein- und Ausgaben gleichzeitig verarbeiten können.

- Interaktivität:

Eine weitere Anforderung an die programmierbaren Bausteine war, dass sie nicht nur autonom arbeiten, sondern auch untereinander kommunizieren können sollten, um so eine Interaktion mehrerer Maschinen zu ermöglichen.

Als Ergebnis wurde unter anderem der Model 120 (Abb. 5) Programmable Brick (1994) vorgestellt, basierend auf einem Motorola Prozessor mit 32 kbyte RAM.

Der Baustein konnte vier Ausgaben und acht Eingaben gleichzeitig kontrollieren, war mit Mikrophon, Infrarot und einer kleinen Anzeige ausgestattet. Programme in Logo-Block (Die zugehörige Programmiersprache [18]) konnten auf den Brick übertragen werden, der sie dann Mittels eines eigenen Logo Interpreters ausführen konnte. So war der P-Brick frei von störenden Kabeln, mit denen beispielsweise noch die Floor Turtle zu kämpfen hatte, da der P-Brick mit einem eigenen Mikrocomputer ausgestattet war, der in einen handtellergroßen Legobaustein eingebettet wurde. Dadurch konnte er an viele verschiedene Lego Kreaturen angesteckt werden und den Konstruktionen der Kinder damit die gewünschten Eigenschaften geben (Abb. 6).



Abb.5 Model 120 [22]



Abb.6 Cricket- Dino [5]



Abb.7 RXC Baustein [22]

Die Funktionsweise des P-Bricks wurde von dem Sponsor LEGO als RCX Lego Brick (Abb. 7) im Rahmen des Lego Mindstorms Construction Kits [8] auf den Markt gebracht. Mit dem enormen kommerziellen Erfolg hatte keiner gerechnet, so wurde in den USA das Lego Mindstorms Robotic Invention Kit allein in den ersten drei Mona-

ten 80.000mal verkauft. Das Set bestand insgesamt aus 717 Teilen, mit LegoBlocks (der zugehörigen Programmiersprache), Motoren, Rädern, verschiedenen Sensoren und den RCX Brick.

2.5 Crickets – die mini Bausteine

1996 entwickelten Martin, Silverman und Berg am Media Lab die neue Generation der programmierbaren Bausteine, die Crickets.

Angelehnt an der Technologie der Thinking Tags [9], die zum 10 jährigen Geburtstag des Lab entwickelt wurden. Die Tags konnten als Namensschilder getragen werden, über Infrarot miteinander kommunizieren und Daten austauschen, gemeinsame oder fehlende Interessen wurden über grüne und rote LEDs angezeigt.

Schließlich stellte sich heraus, dass sich die Kern Hardware der Thinking Tags auch für die programmierbaren Bausteine anbot. Waren diese bisweilen noch so groß wie eine Milchtüte, konnte durch die Tag Technologie die Größe der CPU drastisch verringert werden.

Durch diese Technologie wurde es nun möglich die P- Bricks auf die Größe einer 9 Volt Batterie zu reduzieren, die Crickets [7]. Diese neue Generation der Programmierbaren Bausteine war wesentlich leichter, konnte Motoren kontrollieren, Sensor Daten verarbeiten und via Infrarot mit anderen Crickets und technischen Geräten kommunizieren. Der erste Cricket der im März 1996 am M.I.T. vorgestellt wurde hatte folgende Eigenschaften:

Infrarot Schnittstelle zur Inter-Cricket Kommunikation, sowie dem Herunterladen von Programmen, jeweils zwei Anschlüsse für Ein- und Ausgabe, einen Schalter um das Cricket Programm zu starten und zu stoppen, drei LEDs und eine 9V Batterie zur Energieversorgung.

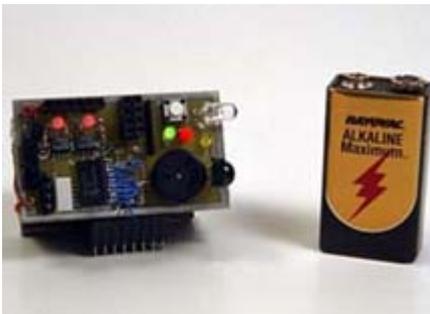


Abb.8 Classic Cricket [10]

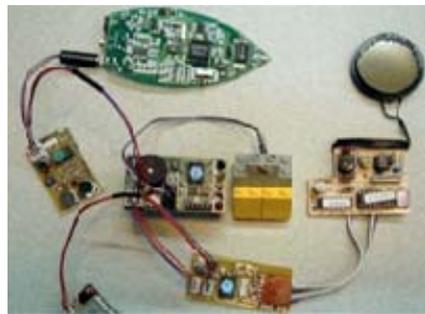


Abb.9 Cricket über einen Bus Gerät verbunden mit anderen Geräten [12]

Bei dem obigen Cricket handelt es sich um den Classic Cricket (Abb.8), um das System aber so klein wie möglich zu halten, begann man Funktionalität auszulagern und in andere Crickets zu integrieren, so entstanden weitere Cricket Arten, wie:

- Display Cricket für Sensor Input Anzeigen
- MIDI Cricket für Sound Input Verarbeitung
- Scientific Cricket für die Analog- Digital Wandlung sonstiger Sensor Eingabe

Um eine Vielzahl an Geräten in das System integrieren zu können, wurde ein eigenes Cricket Bus System entwickelt. Bis zu diesem Zeitpunkt hatte man für jede Funktion einen eigenen Cricket entwickelt, doch mit der Arbeit an dem Bus wurde klar, dass dies gar nicht nötig war. Sie begannen die Schaltkreise zu bündeln und ein eigenes Busgerät zu entwickeln, das mit den einzelnen Geräten (Sensoren, Motoren...) direkt kommunizieren konnte und über ein Protokoll die Informationen an den Cricket weiterleitete. Damit wurden weitere Cricket Versionen überflüssig, da es dem „Blue Dot“ Cricket Bus Gerät (Abb.9) möglich war mit fast jedem anderem Gerät zu kommunizieren.

Programmiert wurden die Crickets in einem Logo Dialekt, der direkt via Infrarot geladen werden konnte.

Aufgrund der geringen Größe der neuen Bausteine entstand ein vielfältiges Einsatzfeld für die Crickets, wie z.B. Roboterprogrammierung, Bewegungserfassung oder interaktive Räume [21]. Da es den Crickets auch möglich war, untereinander zu kommunizieren wurde die Simulation vom Zusammenleben mehrerer Maschinen möglich.

2.6 Das „Beyond Black Boxes“ Projekt - Greifbare Programme von Kindern entwickelt

Im Zuge der Cricket Entwicklung startete das M.I.T. das Beyond Black Boxes Projekt [13,11]. Das Projekt ging über zwei Jahre, den Augenmerk gerichtet auf die Evaluierung und Weiterentwicklung technischer Geräte die Kindern und Jugendlichen erlauben sollten ihre eigenen wissenschaftlichen Instrumente zu bauen. Die geringe Größe der Crickets ermöglichte eine Vielzahl an neuen Konstruktionen, man konnte sie beispielsweise in jeden Gebrauchsgegenstand integrieren oder am Körper tragen. Der geringe Preis (unter 30\$) erlaubte es zudem mehrere Crickets in ein System einzubauen.

Im Folgenden werden einige Beispiele, die aus dem Beyond Black Boxes Projekt hervorgegangen sind präsentiert:

Das Vogelhaus:

Jenny, 11 Jahre entschied sich in dem Workshop eine Maschine zum Vogelfüttern zu entwickeln, dabei ging es ihr hauptsächlich darum eine Möglichkeit zu finden die Tiere zu beobachten und zu analysieren, auch wenn sie



Abb.10 Jennys Vogelhaus [11]

nicht daheim war. Das Grundkonstrukt war schnell gebaut, doch Jenny wollte ja etwas über die Vögel erfahren.

So rüstete sie ihr Vogelhaus mit einem Touchsensor aus der prüfen konnte ob ein Vogel gelandet war. Diesen Sensor verband Jenny mit einem Cricket den sie programmierte die Vögel zu zählen und der einen Lego Mechanismus in Gang setzte welcher eine Kamera auslöste, und somit ein Foto von dem Vogel machte (Abb.10). Nun war es Jenny möglich herauszufinden welche Vögel in ihrer Umgebung lebten. Um zu verhindern, dass derselbe Vogel mehrmals hintereinander fotografiert wurde baute sie in ihr Cricket Programm zusätzlich eine wait- Routine ein die dies verhinderte.

Dancing Creatures:

Durch die Möglichkeit der Crickets miteinander zu kommunizieren, konnten auch Maschinen gebaut werden, die sich „unterhalten“ konnten. Wie die Dancing Creatures (Abb.11), die sich gegenseitig neue Tanzschritte beibrachten, indem sie über ihre Infrarot Verbindung die Programme auf andere Crickets übertragen konnten. (siehe auch 3.3 Dance Craze Buggies)



Abb.12 Dance Craze Buggies [11]

Das Nagelstudio:

Elise (11) entschied sich dafür ihr eigenes Nagelstudio (Abb. 12) zu entwerfen.

Die Funktionalität bestand aus folgendem:

- Eine Vorrichtung in die man die Hände legen konnte
- Sensoren die registrierten ob eine Hand da war
- Ein Motor der Wattebälle bewegte, die die Nägel polierten
- Eine rotierende Feder die die Nägel trocknen sollte
- Hintergrundmusik, solange die Nägel gemacht wurden
- Ein Cricket, der registrierte, wenn eine Hand den Sensor berührte und dann die programmierte Routine zum polieren der Nägel anstieß



Abb.12 Das Cricket Nagelstudio [12]

Durch die hohe Flexibilität des Crickets war ein breites Spektrum an Einsatzmöglichkeiten gegeben, das den Kindern die Möglichkeit gab ihre eigenen Ideen frei zu entwickeln und damit aus der Rolle des reinen Manipulators eines Systems heraus zu brechen. Die kleinen programmierbaren Bausteine konnten nach Belieben in nahezu alle Alltagsgegenstände integriert werden. So wurde mit den Crickets eine Komponente entwickelt, welche die Greifbare Programmierung ein großes Stück nach vorne brachte, da sie durch den hohen Grad der Manipulationsfreiheit des Nutzers die Beziehung zwischen ihm und dem Gerät intensivierte.

Das Beyond Black Boxes (BBB) Projekt hat gezeigt das Kinder aktiv wissenschaftlich arbeiten können mit Hilfe von programmierbaren Bausteinen. Auch wenn die Teilnehmer den technischen Teil nur bedingt verstehen konnten, so haben sie doch schnell gelernt mit diesem kleinen Gerät den statischen Objekten in der realen Welt Verhalten beizubringen.

So bringt der Touchsensor im Vogelhaus alleine noch nichts, aber mit Hilfe des Sensors in Verbindung mit dem Cricket konnte Jenny ein Programm schreiben, das dem Boden Verhalten zuordnete.

Während das BBB Team immer mehr und kleinere intelligente Devices baute, stellte sich schnell heraus wie wichtig die Kommunikation der Geräte untereinander ist damit sie nicht zu technischen Inseln verkommen. So wird als neues Designziel festgelegt, das die Crickets mit möglichst vielen Geräten aus dem Alltag der Kinder (Kameras, Gameboy, Stereoanlage u.v.m.) agieren sollen, damit sie als Teil eines Gesamtsystem gesehen werden, und nicht als isoliertes Spezialgerät.

3 Tangible Programming Brick System – Programmieren nach Baukastenprinzip

3.1. Idee des Systems

Ende der Neunziger entwickelte Timothy McNerney in Anlehnung an die Cricket Architektur das Tangible Programming Brick System[14].

Idee des Systems war einen ganzen Satz Legosteine ähnlicher programmierbarer Bausteine zu realisieren, mit denen die Kinder herkömmlich spielen können und dabei eigenständig Programme entwickeln. Auch McNerney wollte greifbare direkte Programmierung in das reale Umfeld der Kinder bringen:

„My primary motivation is to make programming an activity that is accessible to the hands and minds of younger children by making it more direct and less abstract.” [14]

Während die Crickets über einen Bus mit anderen Komponenten kommunizierten, waren die Tangible Programming Bricks (TPB) mit einer Steckverbindung ausgestattet (Abb.13), wobei das Programm sequentiell von unten nach oben abgearbeitet wurde.

Eine weitere Neuerung fand sich in McNerneys Steckkarten Prinzip (Abb.14), das es ermöglichte den TPB Parameter mitzugeben.



Abb.13 TPB Stapel [14]



Abb.14 TPB mit Steckkarten [14]

Neben der Intention abstrakte Programmierung auf ein kinderfreundliches Level zu bringen sah McNerney auch den großen Vorteil seines Systems darin, dass es mehreren Personen gleichzeitig möglich war an einem Programm zu arbeiten, was nicht nur die Qualität des Programms verbesserte, sondern auch den Lerneffekt innerhalb der Gruppe.

„I imagined a new generation of young programmers building programs together with their hands instead of one person typing on a keyboard and others looking over their shoulder.“ (McNerney) [14]

McNerneys Ziel war es durch die TPB das Programmieren zu einer Handlung zu machen die man anfassen kann, und durch diese Wandlung der Programmierung zu einem Spielzeug, den Kindern fundamentale Konzepte des digitalen Zeitalters in einer spielerischen Umgebung näher zu bringen.

3.2 Die Entwicklung

Der Folgende Abschnitt soll darstellen welche Entwurfsentscheidungen bei der Entwicklung des Tangible Programming Brick Systems zu treffen waren.

[15] McNerneys erster Prototyp (Abb. 15) basierte auf einem 4x2 Legostein. Er enthielt einen Mikrocomputer, eine LED und besaß die Möglichkeit mit seinen direkten Nachbarn nach oben oder unten zu kommunizieren. Doch bald stellte sich heraus, dass die angeforderte Funktionalität in dem kleinen Legostein nicht umsetzbar war, so besaß der Tangible Programming Brick noch keine ausgereifte Kommunikation und keine Eingabemöglichkeiten für Parameter.

Um das System ausbauen zu können stieg McNerney auf den größeren 6x2 Legostein um. Das 6x2 Model (Abb.16) bestand aus einem Mikrocomputer, weiblichen und männlichen Steckverbindungen auf Ober- und Unterseite des Bausteins, wobei die männlichen mit acht vergoldeten Kontakten ausgestattet waren für die Energieversorgung, die serielle Kommunikation und einer globalen run / stopp Signalweiterleitung.

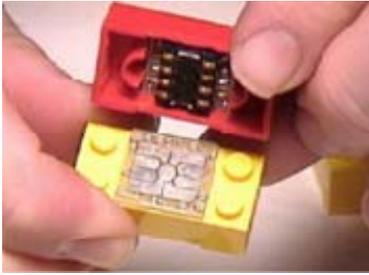


Abb.15 Der 4x2 TPB Prototyp [17]



Abb.16 Der 6x2 TPB [14]

Um den Einsatz der TPB nicht einzuschränken musste bedacht werden, dass die Bricks auch um 180 Grad verdreht (vorne auf hinten) gesteckt funktionsfähig sein sollten, so wurden die Wege für Kommunikation und Energieübertragung doppelt auf einen Baustein angebracht.

Als neue Manipulationsmöglichkeit greifbarer Programme führte McNerney drei Arten von Steckkarten ein, die auf die Vorderseite jedes TPB (Abb.17) passten [14]:

- IR/ Bus Card:
Diese Steckkarte enthielt jeweils einen infrarot Sender und Empfänger, sowie zwei Cricket Bus Konnektoren, durch die der Brick in Cricketsysteme eingebunden werden konnte. Die Infrarotverbindung war bestimmt für das Herunterladen von Programmen und der Kommunikation zu anderen Tangible Programming Brick Systemen.
- EEPROM Card:
Dient zur persistenten Speicherung von Informationen, sowie dem Bereitstellen von Parameterwerten.
- Display Card:
Diese Karte ist in der Lage jegliche alphanumerische Information für den Nutzer anzuzeigen, z.B. Parameternamen, Sensorwerte.

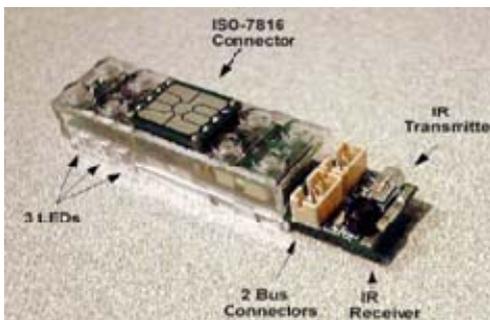


Abb.17
Ein TPB mit einer IR/ Bus Karte [14]

Die Karten erhöhten die Flexibilität der TPB und erlaubten es mehr Funktionalität in das System zu bringen ohne größeren Platz zu beanspruchen.

3.3 Beispiel für greifbare Programmierung mit TPB

Nach der Theorie soll nun ein praktisches Beispiel zum weiteren Verständnis dieses Systems beitragen:

Dance Craze Buggies:

1998 entwickelten Fred Martin und Rick Bovoroy die Anwendung Dance Craze Buggies (Abb.18) für die Crickets, hier konnten Spielzeugautos Tanzschritte beigebracht und unter den Autos weitergegeben werden. Borovoy schrieb über die Idee:



Abb.18

Dance Craze Buggies, die über Crickets kommunizieren können, und über den TPB gefütterten Phaser programmiert werden [16]

„Imagine a child teaching her robotic toy a new dance step. Then, when she is playing with a friend, her toy can „teach“ this new dance to her friend’s toy. Later, her friend can modify this dance a little, and pass it to another friend. The creator of the dance can check the Net to see how far her dance has spread.” [14]

McNerney sah seine Bricks als Möglichkeit die Idee der Dance Craze Buggies zu komplettieren, denn in Borovoys und Martins Szenario fehlte eine geeignete Möglichkeit den Autos neue Tanzschritte beizubringen, ohne dabei am PC programmieren zu müssen. Zur Lösung dieses Problems entwickelte McNerney ein „Lehrer Hilfsmittel“, genannt den Phaser, der mit einem Stapel TPB gefüttert werden konnte[14].

Jedem Brick wurde ein Tanzschritt zugeordnet (vor, zurück, kleiner Schritt vor...), und eine Parameter Karte, die die Anzahl an Wiederholungen festlegte. Dadurch wurde es den Kindern möglich ihre Spielzeugautos ganz ohne PC zu programmieren, indem sie ein entsprechendes Programm mit den Tangible Programming Bricks zusammensteckten und dann den Phaser damit fütterten. Wenn der Phaser aktiviert

wurde, wurde das Programm auf das entsprechende Auto geladen, es hatte den Tanz gelernt.

Diese Idee präsentierte McNerney auf der „Toys of Tomorrow“ Ausstellung 1999. Der Erfolg auf der Ausstellung hat die Idee der Greifbaren Programmierung untermauert, denn nun sind das Programm und das zugehörige Gerät nicht nur logisch verbunden, sondern auch visuell gekoppelt, was es Kindern wesentlich erleichtert das System zu verstehen, im Gegensatz zu einer unsichtbaren Software die vom Computer aus ein Gerät steuert.

Im Moment existieren bereits 50 verschiedenen Tangible Programming Bricks, allerdings wird wohl erst eine Kostenreduzierung der Bausteine dieses System zur Marktreife bringen.

3.4 Wie geht es weiter mit den Tangible Programming Bricks

Die erste Generation der TPB ist entwickelt, und hat ein großes Stück dazu beigetragen die Programmierung greifbar und damit verständlicher zu machen.

Doch auch dieses System steckt noch in den Kinderschuhen, hier noch ein paar Kritikpunkte die McNerney selbst aufgezeigt hat und in Zukunft bearbeiten will [14].

Die TPBs sind begrenzt auf einen linearen Verlauf:

Das jetzige System erlaubt es nur die Bausteine eindimensional zusammenzufügen. Die Implementierung eines Bausteins, der die Verbindung zu Nachbarn zwei- oder dreidimensional erlaubt wäre ein weiterer Entwicklungsschritt. Auch eine „Brücken“ Karte würde das Zusammenfügen in eine zweite Dimension erlauben. Diese Karte könnte als Verbindungsstück zwischen den Dimensionen agieren, womit komplexere Programmabläufe denkbar wären. Was allerdings auch zu einer schwereren Bedienbarkeit dieses Systems führen würde.

Kein visueller Unterschied der TPBs:

Formen können eine große semantische Hilfestellung bieten, gerade bei Kindern fällt das Differenzieren von Objekten leichter durch Formen als durch Labels.

Wie bei einem Puzzle wäre es auch denkbar, das nur bestimmte Bricks zusammenpassen, was eine zusätzliche Implementationshilfe leisten würde.

Reduzierung der Kosten:

Wie bereits erwähnt ist das System im jetzigen Stadium zu teuer für eine Kommerzialisierung, auch wenn McNerney einen großen Kostenfaktor damit ausgeschlossen hat indem er nicht jedem Brick seine eigene Energieversorgung mitgab, so bleibt doch das Verbindungssystem und der Mikroprozessor ein erheblicher Kostenfaktor, wobei sich wahrscheinlich letzteres Problem mit der Zeit von selbst löst.

4 Materielles Programmieren, nur eine Spielerei?

Es gibt weit mehr Entwicklungen in dem Bereich der Greifbaren Programmierung, als ich an dieser Stelle vorstellen kann. Doch ich denke durch das Herausgreifen einiger repräsentativer Beispiele konnte gezeigt werden, das sich viel bewegt in diesem Gebiet, angefangen vor 40 Jahren mit der noch etwas behäbigen Floor Turtle bis zum Ende des 19. Jahrhunderts, in dem es schon möglich ist, den gesamten Rechenprozess in einen Legostein zu packen. Man kann also gespannt sein wohin sich dieser Zweig der Informatik entwickelt.

Doch abgesehen von dem technischen Fortschritt der hier zu ersehen ist, wird auch immer mehr ein Umdenken in der Informatik erkennbar, hin zu mehr Nutzerfreundlichkeit. Das Beispiel der Greifbaren Programmierung zeigt, welche Wege man nehmen kann, um den Menschen den Umgang mit Bits und Bytes zu erleichtern. Manchmal muss man versuchen völlig neue Wege einzuschlagen und alte Denkweisen gänzlich hinter sich lassen, wie hier geschehen indem man nach neuen Interaktionsmöglichkeiten sucht, abseits der Desktop Programmierung. Auch sollte man dieses Thema interdisziplinär betrachten, so nutzt die Technologie alleine nicht viel, es muss auch tief greifend Überdacht werden, was Kinder lernen können und sollen.

Doch es wird auch Kritik laut, die davor warnt, Greifbare Programmierung schon so früh in das Leben unserer Kinder zu bringen wo wir doch jetzt schon immer abhängiger werden von Technologien die die meisten nicht verstehen. Können dann Kinder solche Zusammenhänge überhaupt begreifen, werden sie nicht schon zu früh überfordert? Andererseits ist es enorm wichtig, dass die künftigen Generationen mit dieser Technik vertraut werden, lernen sie zu kontrollieren und effektiv zu nutzen um unsere Zukunft gestalten zu können.

Solange die Kinder spielerisch an das Thema herangeführt werden, und wie in vielen Beispielen gesehen mit großem Einsatz an Greifbaren Programmen arbeiten, denke ich, dass sie der künftigen Generationen die Möglichkeit bieten durch neue Lehr- und Lernmethoden, natürlich digitale Technik in ihren Alltag zu integrieren.

Literatur

1. From Turtles to Tangible Programming Bricks: explorations in physical language Design, Timothy S.McNerney. Per. Ubiquit. Comput. (2004) 8: 326- 337; <http://portal.acm.org/citation.cfm?id=1023813.1023817>
2. Computer Science Logo Style, vol.1, Brain Harvey, Symbolic Computing 2/e Copyright (C) 1997 MIT; <http://www.cs.berkeley.edu/~bh/v1ch10/turtle.html>
3. <http://el.media.mit.edu/logo-foundation/logo/turtle.html>
4. http://www.staff.ucsm.ac.uk/rpotter/ict/logo-control/Screen_Turtle_Notes.pdf
5. Programmable Bricks: Toys to think with, M.Resnick, F.Martin, R.Sargant, B.Silverman, IBM Systems Journal, vol.35, nos.3&4, pp.443-452; <http://researchweb.watson.ibm.com/journal/sj/mit/sectionc/martin.html>

6. <http://web.media.mit.edu/~papert/> (Foto: Seymour Papert)
7. The Children´s Machines: Handheld and Wearable Computers Too, B.Mikhak, F.Martin, M.Resnick, R.Berg, B.Silverman, Proceedings of the International Symposium on Handheld Ubiquitous Computing, Karlsruhe, Germany; <http://www.wellesley.edu/Physics/Rberg/papers/HUC99.pdf>
8. Technologies for Lifelong Kindergarten, M.Resnick, Published in Educational Technology Research & Development, vol. 46, no. 4 (1998); <http://web.media.mit.edu/~mres/papers/lifelongk/lifelongk.pdf>
9. <http://web.mit.edu/6.933/www/Fall2000/LegoMindstorms.pdf>
10. <http://ilk.media.mit.edu/projects/cricket/about/techtalk/reddit.html> (Foto: Cricket)
11. Beyond Black Boxes: Bringing Transparency and Aesthetics Back to Scientific Investigation, Mitchel Resnick, Journal of the Learning Sciences, vol.9, no1, pp.7-30; <http://ilk.media.mit.edu/papers/archive/bbb/>
12. To Mindstorms and Beyond: Evolution of a Construction Kit for Magical Machines, B.Mikhak, F.Martin, M.Resnick, R.Berg, B.Silverman; Morgan Kaufman/Academic Press, San Francisco, CA; <http://www.wellesley.edu/Physics/Rberg/papers/magical-machines.pdf>
13. <http://ilk.media.mit.edu/projects/bbb/>
14. Tangible Programming Bricks: An approach to making programming accessible to everyone, Timothy Scott McNERney, S.M.Thesis <http://xenia.media.mit.edu/~mcnerney/mcnerney-sm-thesis.pdf>
15. Tangible Computation Bricks – Building-blocks for physical Microworlds, Timothy S. McNERney. <http://xenia.media.mit.edu/~mcnerney/tangible-comp-bricks-review2.2.pdf>
16. Meta-Cricket: A designer`s kit for making computational devices, F.Martin, M.Resnick, R.Berg, B.Silverman, IBM Systems Journal, vol.39, nos. 3&4, pp.795-815; <http://www.research.ibm.com/journal/sj/393/part2/martin.pdf>
17. <http://xenia.media.mit.edu/~mcnerney/lego/> (Foto: Der 4x2 TPB Prototyp)
18. LogoBlocks – a graphical programming language for interacting with the world, A.Begel, S.B. Thesis; <http://web.media.mit.edu/~abegel/begelaup.pdf>
19. Logo Philosophy and Implementation, Seymour Papert, Logo Computersystems Inc. 1999; <http://www.microworlds.com/company/philosophy.pdf>
20. Lego/Logo: Learning Through and About Design, M. Resnick, S. Ocko, Norwood , NJ: Ablex Publishing <http://ilk.media.mit.edu/papers/archive/l1.html>
21. Physical Programming: Designing Tools for Children to Create Physical Interactive Environments, J. Montemayor, A. Druin, A. Farber, S. Simms, W. Churaman, A. Dàmour; <ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/2001-21html/2001-21.pdf>
22. <http://el.media.mit.edu/logo-foundation/pubs/logoupdate/v7n1/v7n1-pbrick.html> (Foto: Model 120 Programmable Brick, RXC Baustein)
23. www.papert.org. (Webseite: Seymour Papert, Gründer der Logo- Philosophie)
24. <http://www.st-bartholomews.leics.sch.uk/subjects/ict.html> (Foto: Floor Turtle)
25. <http://pocket.free.fr/html/vm2003/51.jpg> (Foto: Floor Turtle)
26. http://www.bridgendgfl.net/Teachers/digital_portfolio/Images/Year%204/U4Eb.jpg (Foto: Screen Turtle)

Sensorbasierte Benutzerschnittstellen

Nora Zelhofer

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
nora.zelhofer@web.de

Zusammenfassung Sensoren sind in Haushaltsgeräten, in der Gebäudesicherung und in zahlreichen weiteren Anwendungen des täglichen Lebens zu finden. Sensorbasierte Benutzerschnittstellen sind Schnittstellen zwischen Mensch und Maschine, die Sensoren verwenden. Die Sensoren nehmen Veränderungen in der Umgebung wahr und schicken die ermittelten Signale zur Verarbeitung weiter. Sensorbasierte Benutzerschnittstellen dienen zur Vereinfachung von alltäglichen Situationen. Die Benutzung sensorbasierter Gegenstände reicht von der Personenidentifikation in Firmen bis zur Bergung von Lawinenofern. Diese Arbeit soll einen Überblick über die Eigenschaft von Sensoren und deren Einsatzgebiete liefern. Weiterhin wird anhand von vier Beispielanwendungen aus Forschungsprojekten der Einsatz von Sensoren für Benutzerschnittstellen erörtert. Die Arbeit schließt mit einer Diskussion und Zusammenfassung des Themas.

1 Einleitung

Die Menschen sind in ihren Fähigkeiten begrenzt, doch können sie ihre biologischen Sinne für spezielle Aufgaben technisch ergänzen. Mit Sensoren gibt der Mensch auch seinen Maschinen Augen, Ohren und mehr. Dieses Fenster zur Welt ist bei Maschinen der Sensor.

Allein im Alltag begegnen wir Sensoren zu jeder Zeit. Der erste Kontakt mit Sensoren kann z.B. beim Frühstück sein, wenn wir eine Kaffeemaschine mit Temperatursensoren betätigen oder auf der intelligenten Herdplatte etwas zubereiten, auf der nicht einmal die Milch überkocht. Auf dem Weg in die Arbeit mit dem Auto wirken im Hintergrund eine Vielzahl an Sensoren, die wir explizit gar nicht mehr wahrnehmen. Diese verrichten ihre Dienste beispielsweise im Airbag, im Katalysator, beim Motormanagement und in vielen anderen Modulen. Gemessene Größen sind hierbei unter anderem Druck, Gaskonzentration und Beschleunigung. Am Abend, wenn wir den Wetterbericht hören, bekommen wir die Prognosen über das zukünftige Wetter zwar von einem Meteorologen vorgetragen, diese Vorhersage wird jedoch aus Daten, die von Sensoren gemessen werden, errechnet. Diese Daten sind vor Allem Luftdruck, Luftfeuchtigkeit und Lufttemperatur.

„Sensor“ kommt von dem lateinischen Wort „sensus“, ist aber selbst lateinisiert und kann als „Fühler“ übersetzt werden. Ein Sensor erfasst neben chemischen oder

physikalischen Eigenschaften wie Temperatur, Feuchtigkeit, Druck und Beschleunigung, auch die stoffliche Beschaffenheit seiner Umgebung. Die von ihm erfassten Werte oder Zustände werden in der zugehörigen Applikation verarbeitet.[16] Im Allgemeinen werden Sensoren nicht mehr als reine Messwertaufnehmer gesehen, sondern schon als Kombination aus Mess- und Verarbeitungseinheit.

Heutzutage gibt es eine Vielzahl an Sensoren, wie z.B.:

- Beschleunigungssensoren
- Drucksensoren
- Optische Sensoren
- Berührungssensoren

Schon jetzt sind viele Geräte mit Sensoren ausgestattet. Z.B. benutzen GPS-Geräte Lokalisierungssensoren, um die eigene Position zu ermitteln oder jemanden zu navigieren. Ein weiteres Beispiel sind die integrierten Regensensoren im Auto. Diese schalten den Scheibenwischer automatisch ein, falls sie Regentropfen erkennen.

Weitverbreitet sind Sensoren, die Personen identifizieren, z.B. über die Messung des Fingerabdrucks. Die biometrische Erkennung über den Fingerabdruck ist eine billige und weit verbreitete Methode. Zur Erkennung des Fingerbildes wird dieser vorab mit einem Sensor aufgenommen. Für dieses Verfahren ist die Verwendung mehrerer Sensoren möglich: optische, thermische oder kapazitive Sensoren.

Optische Sensoren nehmen das Bild des Fingers mit Hilfe einer Kamera auf. Thermische Sensoren, welche aus einer Matrix von Einzelsensoren bestehen, ermitteln das Wärmebild des Fingers als Datensatz zur Weiterverarbeitung. Kapazitive Sensoren bestehen aus mehreren Kondensatoren. Hierbei werden durch das Auflegen des Fingers auf das Ablesegerät unterschiedliche Kapazitäten bei den einzelnen Kondensatoren gemessen. Werden diese dann ausgewertet, erhält man das Abbild des Fingers. Das gewonnene Abbild wird mit dem eingespeicherten Bild des Fingers verglichen. Besteht eine Übereinstimmung der beiden Bilder, so wird z.B. eine Tür geöffnet.

Heutzutage werden meist die optischen Sensoren benutzt. Problematisch bei dieser Methode ist jedoch die leichte Fälschbarkeit des Abbilds, welche über den Fingerabdruckscanner erstellt wird. Am sichersten gegenüber Täuschung ist die kaum verbreitete thermische Methode. Hier reicht ein zweidimensionales Abbild eines Fingers nicht aus, um sich etwa Zugang zu einem Raum zu beschaffen. Am leichtesten sind die kapazitiven Sensoren zu überlisten, da diese durch einfache Graphitabdrücke eines Fingers zu täuschen sind.[11]

Es stellt sich also heraus, dass Fingerabdrucksysteme bei Zugangssicherung keine wirkliche Kontrolle bieten. In Kombination mit anderen biometrischen Erkennungsverfahren, wie beispielsweise mit der Netzhauterkennung, sind diese Verfahren jedoch wirkungsvoll einzusetzen.

Der Fortschritt in der Sensorik ermöglicht auch das Messen der Emotionen oder des psychischen Zustands eines Menschen. Mit dieser Erkennung kann die Mensch-Maschine-Kommunikation mehr auf menschliche Bedürfnisse eingestellt werden. Solche Emotionserkennungen sind etwa für Designer von Computerspielen interessant.[23]

2 Sensorklassifikation

Heutzutage gibt es eine große Vielzahl an Sensoren, welche in der Technik verwendet werden. Fortschritte in der Sensortechnologie wie z.B. geringer Energieverbrauch oder kostengünstige Produktion ermöglichen verschiedene Integrationen von Sensoren in Geräten und Vorrichtungen.[3] Es gibt nicht nur eine Vielzahl an Anwendungsmöglichkeiten, sondern auch an Sensoren. Die Frage ist nur, welcher Sensor für die jeweilige Anwendung geeignet ist. Die Sensoren können aus logischer und physikalischer Sichtweise unterteilt werden.[4] Dieser vorgeschlagene Ansatz wird in den folgenden Abschnitten näher erläutert.

2.1 Unterteilung der Sensoren aus logischer Sicht

Sensoren werden nach der Art der zu messenden Daten und nicht nach ihrer speziellen Beschaffenheit unterteilt. Somit werden sechs Dimensionen zur Umgebungserkennung unterschieden.[4]

Die erste Dimension ist die Benutzer-Identität. Identität besitzt eine allgemeine Definition, welche von Differenzierung von Objekten bis zur eigentlichen Identifikation reicht. Diese Dimension wird heute schon benutzt, um beispielsweise Dienstleistungen ohne eigentliche Benutzereingaben zu personalisieren oder an den Benutzer anzupassen. Hierbei kann z.B. die Suche nach Produkten in einem Supermarkt mit Hilfe eines PDAs erleichtert werden, indem der Weg zu dem gesuchten Produkt auf dem PDA-Bildschirm angezeigt wird. Die zweite Dimension ist die Ortung bzw. Lokalisierung. Anwendungen aus dem Gebiet der Verteilten Systeme verwenden hauptsächlich Lokalisierungsinformationen, um Objekte finden zu können. Diese Informationen enthalten sowohl 3D Koordinaten wie auch Ortsbeschreibungen.

Die dritte Dimension, die Aktivität, beschreibt die Tätigkeiten des Benutzers. Die Sensoren können die Tätigkeit des Benutzers erfassen, wie z.B. einfache Bewegungsmuster, oder auch die Arbeit präzise beschreiben, die gerade ausgeführt wird.

Die Objektbenutzung ist die vierte Dimension. Sensoren messen in dieser Dimension die Relationen von Benutzern zu Objekten. Außerdem können sie feststellen, wie der Gegenstand benutzt wird und ob er getragen wird.

In der fünften Dimension geht es um die Gemütsbewegungen und die biologischen Zeichen der Benutzer. Sensoren messen das innere Befinden der Benutzer, wie z.B. Zorn, Freude oder Trauer. Forschungen zu diesem Gebiet befinden sich noch in der Anfangsphase. Erste Ergebnisse sind beim Messen von Herzschlag und Hautleitfähigkeit erzielt worden, um die Vorlieben des Benutzers zu ermitteln.

Die sechste Dimension ist die menschliche Interaktion. Hier erfassen Sensoren zwischenmenschliche Beziehungen. Solche Beziehungen können das einfache Zuhören oder Zuschauen sein, oder auch eine Wechselbeziehung, wie sie bei Diskussionen vorkommt.

2.2 Unterteilung der Sensoren aus physikalischer Sicht

Die Platzierung eines Sensors erfordert genaue Überlegungen. Zum Beispiel kann ein Verkehrsstau aus gewisser Entfernung mittels einer Kamera erkannt werden. Eine andere Möglichkeit zur frühzeitigen Stauererkennung wäre mit Sensoren, die in jedem Auto integriert sind. Somit liefern sie Informationen über den Abstand zum Auto, das vor einem fährt und über dessen Geschwindigkeit. Beide Varianten sind für das Lösen des Problems relevant, haben aber unterschiedliche Nebeneffekte. Die Kamera muss einmal montiert werden und ist dann für jedes Auto anwendbar, aber nur für einen Einsatzort. Jedoch findet die Kamera auch andere Einsatzmöglichkeiten, wie beispielsweise als Hilfe für die Polizei zur Suche von Kriminellen. Die lokale Einrichtung an jedem Auto erfordert die individuelle Montage, jedoch kann der Benutzer frei entscheiden, ob er am System mitwirken möchte, oder nicht. Außerdem ist diese Einrichtung am Auto nicht an einen Ort gebunden, sondern kann überall eingesetzt werden.[4]

Es gibt drei verschiedene Kategorien der Sensorplatzierung. Sensoren, die stationär installiert sind, wie z.B. im Boden oder in Wänden, wo Änderungen der Platzierung nur mit Aufwand geschehen, werden als umgebungsgebunden kategorisiert. Sensoren, die am Menschen angebracht sind, können nur dann benutzt werden, wenn sie aktiv getragen werden. Diese Art von Platzierung ist nicht an den Ort gebunden. An Gegenständen befestigte Sensoren hingegen können umgebungsgebunden sein, wenn sie z.B. an einem Stuhl fixiert sind oder können Orts unabhängig sein, wenn sie an einem Schlüssel montiert sind, und dieser vom Mensch mitgenommen wird. Dieser Unterschied hängt von dem Gegenstand ab, an dem der Sensor befestigt wird.

2.3 Zusammenfassung

Die Einteilung der Sensoren ist sehr wichtig für Entwickler und Designer. Fasst man die im vorherigen Abschnitt vorgestellten Unterteilungen zusammen, nämlich die Aufteilung aus logischer wie auch aus physikalischer Sicht, ergibt sich folgende Tabelle, angepasst aus [4].

Tabelle 1. Übersicht über die Sensoren, die zu verschiedenen Zwecken benutzt werden.

Platzierung	Umgebungsgebunden	Objektgebunden	Menschgebunden	Zusammenwirkung
Identität	Biometrische Erkennung	Träge Sensoren	Audio, Träge Sensoren	Lokalisierungssysteme
Objektbenutzung	Audio, Vision	Träge Sensoren	Audio, Träge Sensoren	Lokalisierungssysteme
Ortung	Ladezellen, Audio	GPS	GPS	Lokalisierungssysteme
Emotion	Vision, Audio	Berührung	Träge Sensoren, Temperatur	_____
Aktivität	SmartBoard, Vision	_____	Träge Sensoren, GPS	_____
Interaktion	Audio, Video	_____	Träge Sensoren	_____

Bezüglich [4], ist aus der Tabelle zu erkennen, dass zur Personenidentifikation die biometrischen Sensoren, wie Fingerabdruck oder Netzhauterkennung gut geeignet sind. Andere Methoden, welche auf Audio- oder Sehvermögen beruhen, liefern Ergebnisse mit geringerer Qualität. Trägersensoren, welche auf Gegenständen oder auf Menschen platziert sind, werden verwendet, um Bewegungen zu ermitteln. Um Information über die Position im Freien zu erfassen, wird GPS verwendet. Es ist weiterhin zu erkennen, dass umgebungsgebundene Sensoren die besten Ergebnisse zur Identitätserkennung liefern. Sensoren, welche direkt am Menschen angebracht sind, messen vor allem Gemütsbewegungen, Emotionen und Aktivitäten.

3 Technik

Sensoren werden nach verschiedenen Kategorien, wie schon im vorherigen Kapitel beschrieben, ausgewählt. Zu den jeweiligen Kategorien gibt es verschiedene Sensoren, welche unterschiedliche Funktionen haben. Diese unterschiedlichen Sensoren sollen in diesem Abschnitt vorgestellt werden.

3.1 Optische Sensoren

Licht- oder optische Sensoren (Photodioden, Farbsensoren, UV-Sensoren, usw.) liefern Informationen über die Lichtintensität, die Dichte, die Reflektion, die Farbtemperatur anhand der Wellenlängen und über die Art des Lichts, ob es Sonnenlicht oder künstliches Licht ist. Lichtsensoren sind sehr kostengünstig aufgrund der geringen Fertigungskosten und dem niedrigen Energieverbrauch.

Eingesetzt werden optischen Sensoren etwa bei Lichtschranken. Hierbei werden Unterbrechungen der Leuchtquelle auf dem Sensor registriert und für weitere Verarbeitungen in elektrische Signale umgewandelt.[1]

3.2 Lokalisierungssensoren

Viele gegenwärtige Anwendungen benötigen Informationen über Position, Lage und Nähe von Benutzern oder von Geräten. Auch die Relation zu der Umgebung ist ein wichtiger Faktor. Im Freien wird das GPS (Global Positioning System) verwendet, um Positionen zu bestimmen. Dieses System wird von Satelliten unterstützt, um den genauen Standort zu erfassen.[10] Innerhalb von Gebäuden werden Lokalisierungssensoren in der Umgebung eingebaut, wie beispielsweise das Active Badge System. Bei diesem System trägt jeder Mitwirkende ein Kennzeichen, womit er sofort lokalisiert werden kann, um z.B. ein Telefongespräch gleich in das richtige Zimmer weiterzuleiten. Näheres zum Active Badge System kann unter [9] nachgelesen werden.

3.3 Berührungssensoren

Berührungssensoren (englisch: touch-sensors) sind in verschiedenen Varianten vorhanden, wie z.B. induktive oder kapazitive Versionen.[15]

Weit verbreitet ist die kapazitive Technologie. Abbildung 1 zeigt die Funktionsweise eines kapazitiven Berührungssensors. Solch ein Sensor besteht aus zwei laminierten Glasschichten, welche von einer gemusterten, metallhaltigen aber transparenten Oxidbeschichtung getrennt werden. Ein elektrostatisches Feld wird an der Oberfläche des Sensors erzeugt. Sobald ein Finger oder ein anderes leitfähiges Medium die Oberfläche berührt, wird das elektrostatische Feld verändert und ein Ereignis registriert. Solche Berührungssensoren werden z.B. an den Türen der S-Bahn oder der Tram-Bahn eingesetzt, welche die Türen öffnen, sobald eine Person die markierte Fläche berührt. Berührungssensoren können den Energieverbrauch erheblich reduzieren, da sie nur aktiv werden, wenn sie direkt berührt werden.

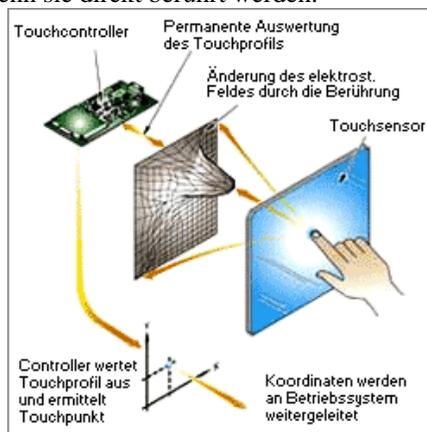


Abb.1: Dieses Bild zeigt die Funktionsweise eines kapazitiven Berührungssensors. Quelle [15]

3.4 Temperatursensoren

Temperatursensoren messen Temperatur und Temperaturänderungen. Dazu wird der fallende bzw. steigende Widerstand im Halbleiter ermittelt. Hierbei ändert sich der Widerstand des Halbleiters in Abhängigkeit von der Temperatur. Es gibt zwei Arten von Temperatursensoren: Thermistoren und Posistoren. Die Unterscheidung geschieht mittels dem Temperaturkoeffizienten, der bei Thermistoren negativ ist, bei Posistoren jedoch positiv. Somit haben Thermistoren einen fallenden Widerstand bei steigender Temperatur, die Posistoren einen steigenden Widerstand bei steigender Temperatur. Mit Thermistoren misst man die Temperatur eines Lebewesens oder der Umgebung. Weitere Anwendungen zur Temperaturmessung mit Thermistoren finden sich in der Haushaltselektronik, in der Heiz- und Klimatechnik und im Auto. Posistoren hingegen dienen als Sicherung bei der Überschreitung vorwählbarer Temperaturen. Einsatzgebiete sind bei der Mess- und Regeltechnik zur Temperaturmessung und als Grenztemperaturfühler beim Motor.[22]

3.5 Schallsensoren

Der Schallsensor ist ein akustischer Sensor, welcher Schwingungen oder Wellen in elektrische Signale umwandelt. In der Regel gehören zur messtechnischen Anwendung ein Sender und ein Empfänger. Schallsensoren werden jedoch nur beim Empfangen gebraucht. Er wird hauptsächlich bei Anwendungen verwendet, welche im Ultraschallbereich arbeiten, also im Frequenzbereich von etwa 40 kHz. Der vom Mensch hörbare Frequenzbereich liegt im Intervall von 15 Hz bis 20 kHz. Diese Wellen sind mit einfachen Mikrofonen messbar. Ultraschallmessungen erfolgen durch die Auswertung der Laufzeitunterschiede zwischen hin- und rückläufiger Welle. Hierbei gilt es jedoch zu beachten, dass Schallwellen von harten Materialien wie Stahl oder Beton reflektiert werden und sich nur in materiellen Medien ausbreiten können.

Im hörbaren Bereich des Menschen werden Schallwandler zur Überwachung von Räumen verwendet, wie z.B. ein Babyphon. Ultraschallsensoren jedoch werden benutzt, um die Dicke undurchdringlicher Medien, wie etwa Stahlplatten, zu bestimmen, sowie Abstände zu messen, wie z.B. bei Einparksystemen beim Auto. Auch zur Bestimmung der Meerestiefe werden Ultraschallwandler, nämlich das Echolot, verwendet.[22]

3.6 Optosensoren

Optosensoren ermöglichen der Technik das „Sehen“. Solch ein Sensor ist auch eine Kamera, welche die Umgebung ablichtet oder filmt und die erhaltenen Informationen weiterleitet.

Sensoren zur Lichterkennung sind Photodioden und Phototransistoren. Sollen Gegenstände von einem Sensor erfasst werden, setzt man selbstauslösende Festkörperbildsensoren ein. Ein anderes, oft verwendetes Prinzip ist das Charged Coupled Device (CCD). Als Maß für den Lichteinfall auf dem Sensor dient die Anzahl der freien Elektronen. Um die gespeicherten Informationen auszulesen, werden die Elektronen periodisch abgegeben. Dieses Prinzip wird beim Scan-Vorgang verwendet.

3.7 Infrarot-Sensoren

Infrarot-Sensoren messen Strahlungen im μm -Bereich. Die Elektronik des Sensors besteht aus einem Photoelement, der die Spannungsänderung misst und diese dann in ein Schaltsignal umwandelt. IR-Sensoren werden zur Kontrolle von temperaturabhängigen Prozessen, wie z.B. in der Nahrungsmittelindustrie, eingesetzt, um die Wärmestrahlung zu messen. Sie reagieren nur auf bestimmte Wellenlängen und sind somit wenig stör anfällig. Infrarot-Sensoren sind auch in Form einer Infrarot-Kamera verfügbar.[22]

3.8 Drucksensoren

Drucksensoren enthalten einen Siliziumchip mit runder oder quadratischer Druckmembran. Sobald auf die Membran ein Druck ausgeübt wird, verformt sie sich entsprechend der Dehnung. Diese Veränderungen zeigen sich in Widerstandsänderungen, die messbar sind. Die gelieferten Ergebnisse werden in elektrische Signale umgewandelt. Wegen des geringen Gewichts der Membran sind Drucksensoren Lage unabhängig. Einsatzgebiete von Drucksensoren sind sehr weitreichend. Sie werden in der Kfz-Elektronik eingesetzt, bei Kühl- und Klimaanlage, bei Getränkeabfüllmaschinen und vielem mehr. Aufgrund ihrer kleinen Größe, den geringen Fertigungskosten und ihrer hohen Genauigkeit sind die Drucksensoren auch in der Medizintechnik einsetzbar z.B. in Beatmungsgeräten oder Blutdruckmessgeräten.[22]

3.9 Beschleunigungssensoren

Beschleunigungssensoren messen die positive und negative Beschleunigung. Beschleunigungsmessungen werden vor allem in den Bereichen der Mechanik und Technik verwendet, um Schaltungen auszulösen oder um Gesten zu erfassen.

Diese Sensoren erfassen Biegung, Positionsveränderung und Deformation des Objekts. Die Beschleunigungskraft lässt sich aufgrund der Auswirkungen auf die Objekte ablesen. Anwendungen, die mit Beschleunigungssensoren arbeiten, dienen entweder der Qualitätssicherung oder der Sicherheit von Personen. Ein Paradebeispiel ist die Airbag-Auslösung bei einem Unfall. Abgesehen von dem Auslösen des Airbags gibt es weitere Anwendungen im Fahrzeug, wie z.B. das Antiblockiersystem. Eine andere Einsatzmöglichkeit von Beschleunigungssensoren ist beim Versand zu finden. Hier können Stöße und somit unsachgemäßer Umgang mit den Paketen gemessen werden. Weitere Anwendungsgebiete sind bei Alarmanlagen, Erdbebenvorhersagen, Vibrationsmessungen an Gebäuden und Maschinen.[1], [22]

3.10 Bewegungsmelder

Ein Bewegungsmelder ist auch ein Sensor, welcher die Bewegungen von Menschen und Tieren in nächster Umgebung messen kann. Diese Art von Sensor wird hauptsächlich zum Ein- oder Ausschalten von Beleuchtungen oder zum Auslösen eines Alarms verwendet. Bewegungsmelder, welche als Lichtschalter eingesetzt werden, enthalten einen Dämmerungsschalter, der die Beleuchtung nur bei Dunkelheit einschalten lässt. Bewegt sich ein Objekt vor dem Melder, so wird das Licht für eine eingestellte Zeitspanne eingeschaltet. Bewegungsmelder, welche als Alarmanlage tätig sind, enthalten keinen Dämmerungsschalter, da sie immer eine Bewegung melden sollen. Wird die Alarmanlage eingeschaltet, löst sie bei jeglicher Bewegung einen Alarm aus.[22]

3.11 Biosensoren

Viele Geräte sind auf Personen bezogen. Biosensoren messen Hautleitfähigkeit, Blutdruck, Atmung und Muskelspannung. Solche Informationen sind bei Anwendungen im sportlichen wie auch im medizinischen Bereich sehr wichtig. Beispielsweise können implantierte Biosensoren bei Diabetikern den Zuckerspiegel messen und bei Bedarf Insulin abgeben. In naher Zukunft wird mit Hilfe von Biosensoren auch das Erkennen des Gemütszustandes einer Person messbar sein.[1]

Es gibt noch viele weitere Arten von Sensoren, deren Auflistung hier zu weit führen würde. Deshalb beschränkt sich diese Arbeit auf die am meisten verwendeten Sensoren.

3.12 Beschränkungen bei der Datenerfassung

Wirtschaftliche und technische Beschränkungen müssen beim Entwickeln von Geräten, welche Zusammenhänge erkennen sollen, betrachtet werden. Technische Einschränkungen entstehen aufgrund des Typs eines Geräts, wirtschaftliche Einsparungen resultieren aus den Kosten eines Gegenstandes.

Die Tragbarkeit und Benutzbarkeit eines Produktes soll nicht auf Kosten der Sensortechnologie verringert werden. Die Größe, das Gewicht und die physikalische Robustheit von Sensoren genauso wie die Integration des Sensors im Gerät sind von großer Bedeutung.

Auch der Energieverbrauch spielt bei batteriebetriebenen Gegenständen, wie z.B. Handys oder PDAs, eine große Rolle.

Um sensorbasierte Geräte erfolversprechend zu bauen, müssen sie sicher und verlässlich sein. Dabei muss das Preis-Leistungs-Verhältnis stimmen. Die zusätzlich eingebauten Sensoren sollen den Geräten nicht zu hohe Kosten und einen zu großen Energieverbrauch schaffen.

Trotz den neu integrierten Sensoren müssen die Geräte bescheiden bleiben, die äußere Erscheinung sollte sich nicht auffällig verändern, um bei Benutzern eine höhere Akzeptanz zu erzielen. Viele Benutzer sind intelligenten Produkten gegenüber skeptisch, denn sie sind um ihre Privatsphäre und um den Datenschutz besorgt.[4]

4 Beispielanwendungen

Schon jetzt arbeiten Forscher an neuen ergonomischen Schnittstellen zwischen Computer und Mensch. In Zukunft sollen normale Tastaturen und Mäuse als Eingabegeräte durch neuere Technologien ersetzt oder erweitert werden.[7] Im Folgenden werden Beispiele vorgestellt, die das Leben des Einzelnen hinsichtlich des Computers oder sogar des Autos erleichtern sollen.

4.1 GestureWrist

Die Gesture Wrist (Gesten-Armbanduhr) ist von Jun Rekimoto, dem Direktor der Interaction Laboratory vom Sony Computer Science Laboratory in Tokio, entwickelt worden.[6] Sie ist erstmals in Zürich auf dem 5. Annual Wearable Computing Symposium vorgestellt worden.

Die Gesture Wrist ist eine Hightech-Uhr, welche dem Benutzer eine Fernsteuerung über tragbare Computer verleiht. In Abbildung 2 ist ein Prototyp der Uhr, die Gesten erkennt, abgebildet. Diese Armbanduhr, welche normal am Handgelenk zu tragen ist, enthält im Armband Sensoren. Die Sensoren erkennen Armbewegungen des Trägers und ob sich die Hand öffnet oder schließt. Die gemessenen Informationen werden an einen tragbaren Computer übermittelt, welcher unmittelbar am Benutzer ist und dessen Bildschirm vielleicht über eine spezielle Brille sichtbar wird. Dieser tragbare Computer kann dann mit dem eigentlichen PC kommunizieren. Somit kann der Benutzer, statt eine PC-Maus zu verwenden, den Mauszeiger auf einem Bildschirm durch vergleichbare Bewegungen mit dem Arm steuern, Icons mittels Öffnen oder Schließen der Hand anklicken.[6], [17]

Die Einzigartigkeit im Design der Gesture Wrist ist die Messung des Handgelenkumfangs, um die unterschiedlichen Handbewegungen zu erfassen. Diese Messung geschieht mit zwei Elektroden, die einander gegenüber am Armband befestigt sind. Somit erkennen die zwei Elektroden die Stellung der Hand und erfassen, ob die Hand geöffnet oder geschlossen ist. Nur wenn sich der Umfang der Hand ändert, weil sich Muskeln anspannen oder Sehnen strecken, wenn ein Finger gezeigt oder eine Faust geballt wird, werden Messungen erzielt, ansonsten werden keine Daten übertragen.

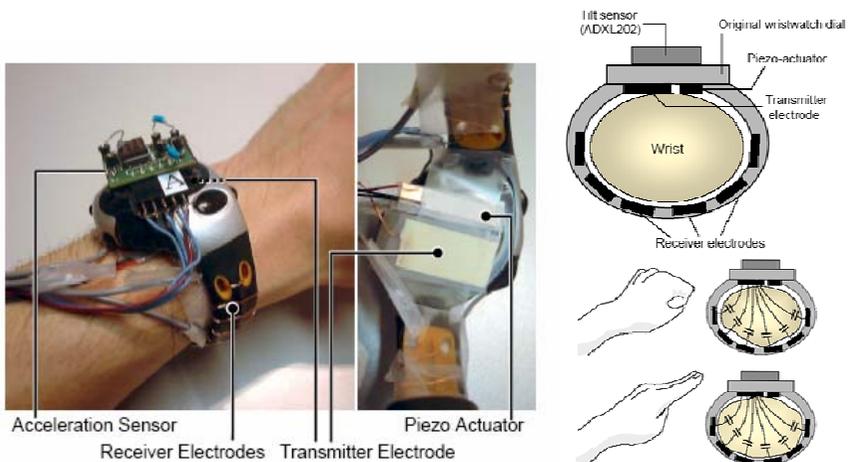


Abb.2: Gesten-Armband als Ersatz für PC-Maus und Messen der Armbewegungen mit Elektroden. Quelle [6]

Weiterhin benutzt die Gesten-Uhr einen Beschleunigungssensor, der direkt neben dem Uhrgehäuse auf dem Armband befestigt ist, um die Richtung und Geschwindigkeit der Armbewegung zu messen.

Zurzeit kann der vorgestellte Prototyp nur einfache Gesten erkennen. Rekimotos System ist noch stark verbesserungsfähig, denn es kann bis jetzt nur identifizieren, ob die Finger ausgestreckt sind und reagiert nur auf einzelne Armbewegungen.

Jedoch ist Rekimoto überzeugt, dass ein feinfühligere System aus Elektroden auch weitere Handbewegungen erkennen könnte.

4.2 ORL Active Floor

Der ORL Active Floor ist ein intelligenter Boden, der von der Olivetty und Oracle Research Laboratory entwickelt wurde.[8] Der Active Floor, der in Abbildung 3 mit zugehöriger Kraftmesszelle abgebildet ist, misst die sich zeitlich verändernde Gewichtsverteilung in einer Umgebung.

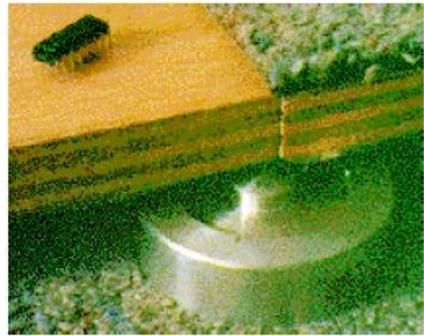
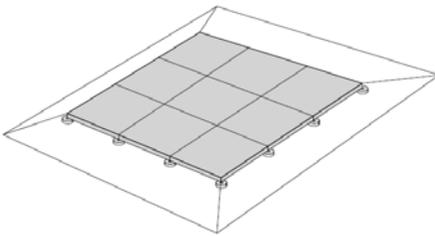


Abb.3: Der intelligente Boden (links) mit zugehöriger Kraftmesszelle (rechts). Quelle [8]

Dieser intelligente Boden ist gegenüber der Gewichtsänderung empfindlich. Die erhaltenen Daten können an das Lokalisierungssystem eines intelligenten Büros weitergeleitet werden. Der Active Floor ist ein quadratisches Raster aus herkömmlichen Teppichfliesen. Diese Teppichfliesen sind mit einer 18 mm dicken Sperrholzplatte und einer 3 mm dicken Stahlplatte verstärkt. Weiterhin sind an jeder Ecke Kraftmesszellen (englisch: load-sensors) angebracht, welche die totale senkrechte Kraft messen. Die verwendeten Kraftmesszellen können einen Gewichtsunterschied von bis zu 50 Gramm bestimmen.

Im Gegensatz zu bestehenden Lokalisierungssystemen, wie etwa das Active Badge System, identifiziert der Active Floor kein Objekt. Es wird lediglich gemessen, dass sich die Gewichtsverteilung auf dem intelligenten Boden verändert hat. Soll dennoch eine Identifizierung stattfinden, so muss nur eine Änderung der Interpretation der Sensordaten erfolgen.

Vorteil dieses Systems ist, dass Objekte und Personen sehr oft in Kontakt mit dem Boden kommen, außer sie sind an der Wand oder Decke befestigt. Weiterhin können Aussagen über die Geschehnisse in der Umgebung gemacht werden.

Der Active Floor erkennt auch, ob ein Utensil beispielshalber vom Tisch entfernt wurde. Diese Möglichkeit der Datenerkennung ist vielleicht in Museen einsetzbar. Weiterhin kann der Active Floor auch als Personenerkennung dienen. Bis jetzt wurde die Identität einer Person anhand der Iris, der Sprache oder des Fingerbildes bestimmt. Mit dem Active Floor kann aufgrund der individuellen Schrittfolge die Person erkannt werden.

Eine aus dem Rahmen fallende Eigenschaft des Active Floor ist das Messen des gesamten Gewichts aller Gegenstände in einem Raum. Wenn beispielsweise drei Personen einen Raum betreten, zeigt der Active Floor eine Zunahme an Gewicht gemäß der Gewichtssumme der drei Individuen, einschließlich den Gegenständen, welche sie an sich tragen. Während eine Person sich bewegt, ändert sich auch die Belastung auf dem Boden, jedoch bleibt das Zentrum der Masse auf der gleiche Höhe, woraus sich das Gewicht der Person erschließen lässt.

Zusammenfassend ist der Active Floor ein Messsystem, welches keine expliziten Tags braucht und welches Objekte in einem klar definierten Raum erkennen kann.

Es gibt viele Anwendungsmöglichkeiten für einen intelligenten Boden. Als Integration in intelligenten Räumen einer Wohnung kann es das Öffnen von Türen, das An- und Ausschalten von Licht und Heizung regulieren oder melden, wenn eine Person gestürzt ist und am Boden liegen bleibt. Mit einem intelligenten Boden in öffentlichen Räumen wie Flughäfen oder Krankenhäuser können Personen ausfindig gemacht werden. Bei Sicherheitssystemen kann der Active Floor eine Erweiterung zu Erkennungssystemen mit Tags sein. Wenn z.B. eine Person, per Tag-Identifikation in einen Raum tritt, wird dieser auch per Gewichtsänderung erkannt. Somit kann aufgrund der Änderung der Belastung des Bodens erfasst werden, ob sich die Person zu nah an laufende Maschinen nähert oder ob ein Gegenstand aufgrund der hohen Gewichtsänderung aus dem Raum entfernt wurde.

Eine andere Alternative zum Active Floor ist ein intelligenter Teppich.[21] Er enthält bis zu 25 Sensoren und Mikrochips pro Quadratmeter, welche auf der Teppichunterseite miteinander vernetzt sind. Die Chips sind als Alarmsystem einsetzbar und unterstützen die Klima- und Leitsystemtechnik in Gebäuden. Drucksensoren registrieren Personen, die sich auf dem Teppich aufhalten und wohin sie gehen. So kann etwa ein Alarm ausgelöst werden, wenn ein Einbrecher durch das Fenster in eine Wohnung steigt.

4.3 Sensoreinsatz bei der Bergung von Lawinopfern

Aufgrund der steigenden Opferzahlen, die durch ausgelöste Lawinen entstehen, entwickelten Forscher ein System zur Bergung und Lokalisierung der Opfer.[5] Statistiken zeigen, dass dreiviertel aller Lawinopfer an Ersticken sterben. Eine erfolgreiche Bergung der Opfer muss innerhalb der ersten fünfzehn Minuten geschehen. Heutzutage gibt es schon Ortungsgeräte, die von den Aktivisten selbst getragen werden. Diese Geräte liefern Informationen über die Richtung und Entfernung des Unfalls. Mit deren Hilfe können schon erste Such- und Rettungsmaßnahmen gestartet werden.

Leider reichen diese zwei Information über die Richtung und Entfernung der verschütteten Opfer nicht immer aus, um sie zu retten. Erweitert man diese bestehenden Ortungsgeräte mit Sensoren, so kann zusätzlich noch über den emotionalen Zustand,

über Lebenszeichendaten und über die Bewegungsmuster oder über die Orientierung der Opfer im Schnee informiert werden. Mit diesen zusätzlichen Mitteilungen können die Menschen schneller und sicherer geborgen werden.

Um die Überlebenschancen von Lawinenofern zu erhöhen, indem sie schneller geborgen werden, hilft der Einsatz von Sensoren, welche in der Kleidung der Opfer integriert sind. In Abbildung 4 wird ein Prototyp mit den zugehörigen Sensoren gezeigt.



Abb.4: Prototyp mit Beschleunigungssensor und Biosensor. Quelle [5]

Eingesetzte Sensoren sind Beschleunigungssensoren und Biosensoren. Beschleunigungssensoren liefern Informationen über die Orientierung der verschütteten Personen. Diese Informationen werden über die Richtung der Schwerkraft ermittelt. Somit können Verletzungen während des Ausgrabens vermieden werden. Mit Biosensoren werden Puls, Bewusstsein und Atmungsaktivität gemessen. Aufgrund dieser Informationen können sich die Sanitäter auf die Behandlung vorbereiten, denn jeder Zustand erfordert andere Herangehensweisen. Zudem erfahren die Sanitäter im Krankenhaus oder im Krankenwagen, wie viele von ihnen gebraucht werden, denn eine Wiederbelebung beansprucht z.B. zwei Helfer.

Die getragenen Ortungsgeräte von den verschütteten Personen müssen dann von den Bergungs-Leuten gefunden werden. Eine Möglichkeit ist per Audio-Unterstützung, so dass das „Finde-Gerät“ lauter piept, je näher es an die Lawinenoferne kommt. Ein anderer Weg wäre per visueller Unterstützung, so dass Pfeile die Richtung visualisieren und auch Abstände zu dem Opfer gezeigt werden.

Folglich bekommen die Helfer und Retter durch den Einsatz an Sensoren wichtige Informationen geliefert, mit deren Hilfe sie lebensrettende Schritte einleiten können.

4.4 Das „sehende Auto“

Sensoren spielen auch eine große Rolle in der Automobilelektronik. Schon heute sind Sensoren in Autos integriert, wie etwa ABS, ESP und Airbag. Das „sehende Auto“ kann man einerseits als Entwicklung zum Autopilot sehen, andererseits als Hilfe, die Verletzungsgefahr bei Unfällen zu mindern.[12], [14]

Beim Ersteren soll das Auto mittels Sensoren das Umfeld wahrnehmen und interpretieren. Somit kann es auf Geschehnisse und Hindernisse reagieren, um das Fahren sicherer und komfortabler zu machen. Ziel dieser Entwicklung ist, die Schwere von Unfällen zu reduzieren und die Anzahl der Unfälle, die aufgrund der Unaufmerksamkeit des Fahrers geschehen, zu verringern.

Beim zweiten Ansatz sollen Airbags mit Hilfe von Sensoren der Situation entsprechend reagieren und Passagiere schützen, falls es doch zum Unfall kommt. In den Autositzen sind Sensormatten eingelassen, welche das Gewicht und die Gewichtsverteilung messen. Folglich erkennt das Auto, ob ein Kind, ein leicht- oder schwergewichtiger Erwachsener oder ein Gegenstand sich auf dem Sitz befindet. Weiterhin sollen Digitalkameras den Innenraum eines Autos überwachen, damit die Positionen der Passagiere erfasst werden können. Wenn eine Person beispielsweise während eines Zusammenstoßes nach vorne gebeugt ist, bläst sich der Airbag langsamer auf, damit Verletzungen durch den Luftsack vermieden werden.

5 Chancen und Grenzen

Prozessoren und kleinste Sensoren werden aufgrund der Miniaturisierung der Computertechnologie in absehbarer Zukunft in Alltagsgegenstände verstärkt integriert werden. Die traditionellen Eingabegeräte wie Tastatur und Maus werden verschwinden und die Kommunikation zu Computern wird über unsere Kleidung, Armbanduhren oder Möbel erfolgen.[2], [19]

Sensoren bieten der heutigen Gesellschaft sowohl Chancen als auch Grenzen. Die Chancen liegen etwa darin, Computerspiele realistischer, Computereingaben effektiver und einfacher zu machen und Alltagsgegenstände wie Geschirrspüler, Kühlschränke, usw. individueller und intelligenter zu gestalten. Auch das Leben des Einzelnen kann sicherer werden, denn bei einem Verkehrsunfall können sie z.B. per GPS schnell geortet und gerettet werden. Ein anderer Aspekt wäre, wenn in Zukunft bei künstlichen Gelenken, wie etwa Hüft- oder Kniegelenken, Mikro-Sensoren integriert sind, die eine Infektion bemerken, die Bakterien bestimmen und diese mittels geeignetem Antibiotikum bekämpfen. Kritiker argumentieren jedoch, dass die Kosten einer solchen Operation finanziell nicht zu tragen sind.[13]

Auch wenn die sensorbasierten Benutzerschnittstellen Fortschritt und Effizienz versprechen, so haben sie dennoch einen negativen Effekt, und zwar rechtlicher und ethischer Art. Das Problem des „gläsernen User“ wird durch die Benutzung von Sensoren vergrößert, da die Informationen vor Dritten nicht sicher sind. Wenn Sensoren in der Kleidung integriert sind, senden sie Daten, ohne dass der Träger dies mitbekommt. Diese Daten können anderweitig ohne das Einverständnis des Trägers weiterverwendet werden.[18]

Somit haben sensorbasierte Benutzerschnittstellen weit reichende Folgen für die Bereiche Sicherheit und Datenschutz, denn eine Überwachung der einzelnen Personen ist ohne Anstrengung möglich. Informationen von Sensoren, welche dauerhaft angelegt werden, können von unberechtigten Dritten gelesen werden und für andere Zwecke missbraucht werden. Emotionserkennung kann etwa in einer Firma benutzt werden, um die Motivation der Angestellten zu überprüfen. Basierend auf diesen Informationen kann ein Angestellter entlassen oder befördert werden. Der Schutz der Privatsphäre des Einzelnen ist gefährdet, wenn sensorbasierte Benutzerschnittstellen immer aktiv sind und regelmäßig Daten versenden, um dem Benutzer jederzeit zur Verfügung stehen zu können. Somit ist die Balance von Freiheit und Sicherheit sehr gefährdet, denn die qualitativen und quantitativen Möglichkeiten der Überwachung werden ausgeweitet zu einer dauerhaften und heimlichen Kontrolle.

6 Schlussfolgerung

Zusammenfassend lässt sich sagen, dass der Einsatz von Sensoren positive wie auch negative Konsequenzen haben kann. In der Umwelt eingesetzte Sensoren können ökologische Effekte besser kontrollieren, aber die Überwachungsmöglichkeit der Menschen könnte schwerwiegende Effekte nach sich ziehen, wie etwa Einschränkungen der Privatsphäre.

Der anhaltende Trend des Fortschritts in der Mikroelektronik führt dazu, dass elektronische Komponenten noch bemerkenswert leistungsfähiger, kleiner und billiger werden. Demzufolge werden sensorbasierte Benutzerschnittstellen in absehbarer Zukunft eine große Rolle in unserem alltäglichen Leben spielen. Sie erleichtern die Kommunikation zwischen Mensch und Maschine und machen unser Leben sicherer, nicht nur durch den Einsatz in Automobilen, sondern auch in Kleidungen zur Lokalisierung. Doch auch in der Medizin werden Sensoren vermehrt eingesetzt. So sollen Atmungsgeräte intelligenter werden oder die Abstoßung von transplantierten Organen schneller erkannt werden. Auch die Behandlung von chronisch kranken Patienten kann verbessert werden, wenn Sensoren kontinuierlich Daten an den behandelnden Arzt senden. Weiter kann auch das Verhalten von autistischen Kinder erforscht werden, wenn Sensoren in das tägliche Leben eines Kindes integriert sind.[13]

Bei der Verwendung von Sensoren spielt nicht nur die Anwendung eine große Rolle, auch die Effizienz und Wirtschaftlichkeit sind wichtige Faktoren. Manche industriellen Branchen wie etwa der Maschinen- oder Automobilbau sind ohne Sensorik nicht mehr denkbar. So werden Sensoren zur Massenware und sichern Arbeitsplätze in der Sensorherstellungstechnik. Technischer Fortschritt, steigende Wirkungsgrade, geringere Emissionen, bessere Ressourcen- und Umweltschonungen können durch den Einsatz von Sensoren ermöglicht werden.

Aber Sensoren sind auch Gegenstand und Mittel der Forschung, denn jeder Sensor basiert auf einem naturwissenschaftlichen Effekt und nutzt den funktionellen Zusammenhang zweier oder mehrerer Größen.

Viele weitere Anwendungsmöglichkeiten intelligenter und kommunizierender Gegenstände sind denkbar. Grundsätzlich beruhen die Grenzen ihres Einsatzes weniger auf technischen Gesichtspunkten, sondern sind eher ökonomischer, rechtlicher und

moralischer Art. Denn wer bestimmt, was die Alltagsgegenstände wem verraten und was sie sich merken dürfen? Auch wenn solch intelligente Produkte für den Endverbraucher noch weitgehend undenkbar sind, dürfte langfristig die Techniken der Verteilten Systeme eine große wirtschaftliche Bedeutung bekommen, da folglich innovative Produkte und neuartige Dienstleistungen möglich werden.

Literatur

1. A. Schmidt and K. V. Laerhoven, "How to build smart appliances," IEEE Personal Communications, August 2001
2. A. Schmidt, M. Strohbach, K. Van Laerhoven, and H.W. Gellersen, "Ubiquitous Interaction - Using Surfaces in Everyday Environments as Pointing Devices," 7th ERCIM Workshop "User Interfaces For All", 23 - 25 October, 2002
3. B. Yoshimi, "On Sensor Frameworks for Pervasive Systems" Workshop on Software Engineering for Wearable and Pervasive, Computing SEWPC00 at the 22nd Int. Conference on Software Engineering ICSE 2000.
4. F. Michahelles, B. Schiele, "Sensing Opportunities for Physical Interaction." Workshop on Physical Interaction (PI03) at Mobile HCI, Udine, Italy, 2003.
5. F. Michahelles, P. Matter, A. Schmidt, B. Schiele., "Applying Wearable Sensors to Avalanche Rescue: First Experiences with a Novel Avalanche Beacon". Computer & Graphics, Vol. 27, No. 6, 2003
6. J. Rekimoto, "GestureWrist and GesturePad: Wearable Interaction Devices," Fifth International Symposium on Wearable Computers, 2001
7. LE. Holmquist, R. Maz, and S. Ljungblad. "Designing Tomorrow's Smart Products - Experience with the Smart-Its Platform." In DUX, San Francisco, 2003
8. M. Addlesee et al., "The ORL Active Floor," IEEE Personal Communications, vol. 4, no. 5, pp. 35-41, October 1997
9. R. Want et al., "The Active Badge Location System," ACM Trans. Information Systems, vol. 10, no. 1, Jan. 1992, pp. 91-102.

Internetadressen, aufgerufen im Zeitraum November, Dezember 2004:

10. http://de.wikipedia.org/wiki/Global_Positioning_System
11. <http://www.bromba.com/faq/fqfaqd.htm>
12. http://www.eit.uni-kl.de/fb_et/news/vortrag_01_07_04.htm
13. <http://www.heise.de/tp/r4/artikel/5/5501/1.html>
14. http://www.iitb.fraunhofer.de/servlet/is/1620/Sehendes_Auto.mpg
15. http://www.infotronik.at/touchscreen.php?id=7&category=touchscreen_komp
16. <http://www.lexikon-definition.de/Sensor.html>
17. <http://www.newscientist.com/article.ns?id=dn1398>
18. <http://www.privacy-security.ch/interface/2004/pdf/Buellesbach.pdf>
19. <http://www.vs.inf.ethz.ch/publ/papers/ethbulletin03.pdf>
20. <http://www.weltdterphysik.de/forschung/experimente/sonstiges/sensoren/>
21. <http://www.zdnet.de/z/news/0,39023863,2134237,00.htm>
22. <http://www.fbp.mfh-iserlohn.de/Labore/ELON/ex296.pdf>

Referenz auf Artikel aus dem Seminar UI-Update 1.0

23. D. Kern, Gefühlsorientierte Benutzerschnittstellen, Seminar UI-Update 1.0, Ludwig-Maximilians-Universität München, Feb. 2005
<http://www.hcilab.org/events/ui-update1.0/>

Kontextbezogene Benutzungsschnittstellen

Oliver Falke

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`falke@ifi.lmu.de`

Zusammenfassung Diese Arbeit ist im Rahmen des Hauptseminars „New User Interfaces“ an der Ludwig-Maximilian-Universität München entstanden und soll eine Einführung in das Gebiet der kontextbezogenen Benutzungsschnittstellen geben. Es werden verschiedene Definitionen von Kontext untersucht und darauf aufbauend wird beleuchtet wie man verschiedene Arten von Kontext feststellen kann. Unterschieden werden dort drei Klassen: der Ort, Kontext auf niedriger und Kontext auf höherer Ebene. Weiter folgt eine Untersuchung der Modellierungsarten für Kontext, d.h. welche Datenstrukturen bieten eine gute Möglichkeit Kontext abzuspeichern und auszutauschen. Eine Untersuchung der Nutzung von Kontext in Benutzungsschnittstellen wird ebenfalls durchgeführt, gefolgt von einem Blick auf die Themen Sicherheit und Privatsphäre, die bei der Betrachtung von kontextbezogenen Systemen oft außen vor gelassen werden. Abschlossen wird die Arbeit durch die Vorstellung mehrerer Systeme, die durch die Verwendung von Kontext einen Zusatznutzen für den Benutzer erreichen.

1 Einleitung

In den letzten Jahren ist die Verfügbarkeit von leistungsfähigen, mobilen Endgeräten drastisch gestiegen. PDAs und Mobiltelefone weisen heutzutage teilweise die Rechenleistung auf, die noch vor wenigen Jahren nur bei stationären PCs gefunden wurde [9]. Des Weiteren konnten z.B. GPS-Empfänger in den letzten Jahren sehr stark miniaturisiert werden, was 2004 zur Einführung des ersten PDA mit integriertem GPS-Empfänger führte.

Parallel dazu hat sich die Verfügbarkeit von schnellen drahtlosen Netzen stark erhöht. So wurden mit GPRS und UMTS Technologien eingeführt, die eine schnelle mobile Datenübertragung ermöglichen und bereits annähernd flächendeckend verfügbar sind, bzw. in absehbarer Zeit verfügbar sein werden. Hinzu kommt die rasante Verbreitung von Wireless-LAN-Hotspots nach dem IEEE 802.11 Standard. Zudem sind die Kosten für mobile Endgeräte und die mobile Datenkommunikation stark gefallen und diese Entwicklung wird sich wohl, zumindest bei der Datenkommunikation, auch in den nächsten Jahren fortsetzen.

Dadurch wandelt sich auch die Art des Umgangs mit Computern. Die Anwender wollen nicht mehr nur an ihrem Arbeitsplatz einen Computer bedienen sondern überall seine Unterstützung nutzen. Außerdem möchten sie heute überall auf ihre gespei-

cherten Daten zugreifen können. Die moderne Gesellschaft bewegt sich immer mehr in Richtung Mark Weisers Vision von ubiquitär verfügbaren Computern [17].

Zudem werden Benutzungsschnittstellen und die darunter liegenden Systeme immer komplexer und unüberschaubarer für den Anwender. Die Flut von Informationen macht eine effiziente Bedienung der Benutzungsschnittstellen immer aufwändiger. Dies Alles sind Voraussetzungen dafür, warum die Verwendung von Kontext in Systemen und Benutzungsschnittstellen immer interessanter wird.

2 Was ist Kontext?

Der Begriff Kontext ist sehr weitläufig und wird in vielerlei Hinsicht verwendet. Im Online Wörterbuch *Merriam-Webster Online Dictionary* lautet die Definition von Kontext „in gegenseitiger Beziehung stehende Konditionen, in denen etwas existiert oder passiert“ [9]. Diese sehr allgemeine Definition hilft uns aber nicht weiter, aus diesem Grund wurde eine Vielzahl weiterer Definitionen für Kontext entworfen. Im Folgenden werden nun einige der gebräuchlichsten vorgestellt.

- Eine der ersten Definitionen stammt von Schilit et al., die den „Ort, die Identitäten von Personen und Objekten in der Umgebung, sowie Änderungen bezüglich dieser Objekte“ als Kontext bezeichnen. [12], [11]
- Brown et al. zählen Elemente von Kontext einfach auf: „der Ort, Identitäten von Personen in der Umgebung des Benutzers, Tageszeit, Jahreszeit, Temperatur, usw.“. [13]
- Ryan et al. definieren Kontext als „den emotionalen Zustand des Benutzers, den Fokus seiner Aufmerksamkeit, Ort und Ausrichtung, Datum und Uhrzeit, Objekte und Personen in der Umgebung des Benutzers“. [14]

Definitionen, die Kontext lediglich auflisten, sind schwierig anwendbar, da unklar ist, ob ein bestimmter Begriff der nicht aufgelistet ist zum Kontext gehört oder nicht [11].

Andere Forscher haben versucht den Begriff Kontext nur als Synonym zu verwenden, beispielsweise für Umgebung oder Situation. Aber selbst hier wird der Begriff sehr unterschiedlich verwendet, einige Forscher bezeichnen die Umgebung des Benutzers als Kontext, andere die Umgebung der Applikation.

Eine spätere Arbeit von Schilit et al. [4] definiert Kontext als „Wo man ist, mit wem man ist und welche Ressourcen in der Nähe sind?“, was wohl eine der anschaulichsten Definitionen darstellt. Kontext ist also die sich ständig verändernde Ausführungsumgebung, die folgende Elemente umfasst [11]:

- Die Computerumgebung: verfügbare Hardware, Ein- und Ausgabegeräte, Netzwerkverbindungen, Kosten der Computernutzung, usw.
- Die Benutzerumgebung: Ort, Menge der Personen in der näheren Umgebung, soziale Situation
- Die physische Umgebung: Lichtverhältnisse, Geräuschpegel, usw.

Dey et al. nehmen Elemente aller bisherigen Definitionen auf und kommen zur folgenden Definition:

„Kontext ist jegliche Information, die verwendet werden kann um die Situation einer Entität zu beschreiben. Eine Entität ist eine Person, ein Ort oder ein Objekt, die für die Interaktion zwischen einem Benutzer und einer Applikation als relevant erachtet werden, inklusive dem Benutzer und der Applikationen selbst“ [11].

Diese Definition lässt sowohl implizite, als auch explizite Informationen zu, das heißt es ist irrelevant ob das System die Information selbst ermittelt, oder der Benutzer diese durch eine Eingabe selbst angibt.

Dies sind nur die gängigsten Definitionen des Begriffs Kontext, es gibt zusätzlich noch eine Vielzahl weiterer. Wie man gesehen hat, basieren viele der Definitionen auf früheren Definitionen anderer Forscher. Schritt für Schritt wurde sie so bis heute erweitert und angepasst.

3 Wie stellt man Kontext fest?

Wie man im vorigen Abschnitt gesehen hat, gibt es eine unüberschaubare Vielzahl an Definitionen von Kontext. Hinzu kommt, dass die Menge der Informationen, die zum Kontext gezählt werden, ebenfalls sehr groß ist. Nun stellt sich die Frage, wie man diesen Kontext ermitteln kann. Die einfachste Möglichkeit besteht darin, dass das System den Anwender danach fragt und der Anwender den Kontext eingibt. Diese Möglichkeit soll hier allerdings nicht mehr weiter untersucht werden, da sie ja die Bedienung der Benutzungsschnittstelle noch weiter erschweren würde. Das System soll den Kontext folglich eigenständig ermitteln, damit der Benutzer nicht mit der zeitaufwendigen Eingabe des Kontexts belastet wird und um damit eine möglichst einfach zu bedienende Benutzungsschnittstelle zu erhalten. Grundsätzlich unterscheidet man zwei verschiedene Vorgehensweisen. Zum einen kann man das Benutzerverhalten bei der Benutzung einer Applikation aufzeichnen und daraus Kontextinformationen gewinnen. Zum anderen kann man die Umgebung oder das mobile Endgerät „intelligent“ machen, beispielsweise mit Sensoren. Im Folgenden untersuchen wir anhand der Art des Kontexts, wie man ihn feststellen kann.

3.1 Ort

Die Kontextinformation Ort ändert sich, sobald der Benutzer seinen Standort verändert. Es gibt bei den verschiedenen Ortungssystemen große Unterschiede hinsichtlich der Genauigkeit der Lokalisierung. Für einige Applikationen mag eine Ortung im 100 Meter-Bereich genügen, wohingegen andere Applikationen eine Auflösung von wenigen Metern oder sogar Dezimetern erfordern. Prinzipiell unterscheidet man bei der Feststellung des Ortes, ob sich der Benutzer im Freien befindet oder innerhalb geschlossener Gebäude.

Um den Kontext außerhalb von Gebäuden festzustellen, bietet sich das Global Positioning System (GPS) an. Dabei erfolgt die Ortung über Satelliten, wobei eine Genauigkeit von etwa 10 bis 20 Metern erreicht wird. Der große Vorteil dieser Technologie ist die weltweite Verfügbarkeit, der von Satelliten ausgestrahlten Or-

tungssignale. Eine Positionsbestimmung über kommerzielle Mobilfunknetze wäre eine Alternative, wobei sich momentan die Genauigkeit der Ortung noch auf einzelne Mobilfunkzellen beschränkt und somit nicht exakt genug für die meisten kontextbezogene Anwendungen ist.

Unglücklicherweise besteht innerhalb von Gebäuden nicht die Möglichkeit der Positionsbestimmung per GPS, da die Signale der Satelliten nicht oder nur unzureichend empfangen werden können. Viele Projekte haben in der Vergangenheit aus diesem Grund eigene Ortungssysteme für ihre Anwendungen entwickelt. Die verwendeten Technologien umfassen Infrarot, Funk und Ultraschall. Die Mehrheit der Ortungssysteme für den Innenbereich arbeitet zellbasiert. Dabei ermittelt entweder das Objekt die Zelle in der es sich befindet, oder das System ortet das Objekt in einer bestimmten Zelle. Andere Verfahren versuchen anhand der Signalstärke von Funksignalen die Distanz zwischen Sender und Empfänger zu ermitteln, und vergleichen die ermittelte Signalstärke mit zuvor gemessenen Referenzwerten in einer Datenbank [9].

Ein ungewöhnlicher Ansatz zur Positionsbestimmung wurde beim GUIDE-Projekt [15] verfolgt. Hier wurden Basisstationen nach dem IEEE 802.11-Standard aufgestellt und deren Funkbereich als Ortungszelle verwendet. Dieser Ansatz hat den Vorteil, dass die Verfügbarkeit sowohl in Gebäuden als auch im Freien gegeben ist, allerdings setzt dies eine Versorgung durch die Basisstationen im gesamten Applikationsgebiet voraus. Dies ist bei Systemen die z.B. eine landesweite Ortung benötigen nicht praktikabel.

Zusammenfassend kann man sagen, dass es bis heute kein System gibt welches in der Lage ist sowohl innerhalb von Gebäuden, als auch im Freien eine Positionsbestimmung mit hoher Genauigkeit durchzuführen und zudem eine hohe Verfügbarkeit aufweist. Aus diesem Grund werden in vielen Projekten unterschiedliche Technologien zur Ortung kombiniert. Viele Systeme verwenden die Infrastruktur, mit der Daten übertragen werden ebenfalls für die Ortung, wohingegen dies bei Systemen die beispielsweise GPS verwenden nicht der Fall ist.

3.2 Kontext auf niedriger Ebene

Neben dem Ort gibt es noch eine Vielzahl weiterer Kontextinformationen die auf niedriger Ebene ermittelt werden können.

- **Zeit:** Diese Kontextinformation kann meistens einfach über die eingebaute Systemuhr ermittelt werden. Neben der Uhrzeit gibt es noch andere Formen dieser Kontextinformation, wie z.B. Tag der Woche, Jahreszeit, Zeitzone. Diese erweiterten Zeitkontextinformationen wurden bisher aber noch in keinen Applikationen verwendet [9].
- **Objekte in der Nähe:** Falls eine Datenbank mit dem Aufenthaltsort von Objekten im System existiert, kann durch einfache Datenbankabfragen die Menge der Objekte in der Nähe ermittelt werden.
- **Netzwerk-Bandbreite:** Dies ist eine sehr wichtige Kontextinformation, welche allerdings schwierig zu ermitteln ist und sich außerdem ständig ändern kann. Meist

wird hierfür Unterstützung vom Basissystem benötigt, beispielsweise durch Funktionen des Betriebssystems.

- **Orientierung:** Die Orientierung eines Gerätes kann einfach durch die Anbringung von zwei rechtwinklig angebrachten Orientierungssensoren ermittelt werden. Dies wird beispielsweise im System von Schmidt et al. [1] verwendet.

Weitere Kontextinformationen wie Lichtverhältnisse, Geräuschpegel, Feuchtigkeitsverhältnisse, Vibration usw. können ebenfalls durch einfache Sensoren ermittelt werden.

3.3 Kontext auf höherer Ebene

Unter Kontext auf höherer Ebene versteht man Kontextinformationen die sich aus komplexen sozialen Informationen zusammensetzen. Zum Beispiel wäre die momentane Aktivität oder Beschäftigung des Anwenders eine solche Information. Dies zu ermitteln ist äußerst schwierig und es gibt unterschiedliche Ansätze dies zu erreichen. Eine Variante wäre den Kalender des Anwenders zu konsultieren. Aber es steht nicht immer im Kalender, was die Person den ganzen Tag zu tun beabsichtigt. Andere Projekte versuchen mittels Videokamera und Bildverarbeitungsalgorithmen die Information zu berechnen. Wieder andere Wissenschaftler versuchen durch Kombination mehrerer Kontextinformationen auf niedrigerer Ebene, die Kontextinformation der höheren Ebene zu ermitteln.

4 Wie modelliert man Kontext?

Wie man in den vorigen Abschnitten gesehen hat, gibt es eine fast unüberschaubare Vielzahl von Kontextinformationen. Zudem können diese Kontextinformationen unterschiedlichster Art sein, so dass es nicht immer einfach ist den Kontext, in für Maschinen geeigneter Form, zu modellieren. Alle bisher entwickelten Systeme benutzen ihre eigenen Verfahren um Kontext zu modellieren. Dies macht es unmöglich, Kontextinformationen zwischen solchen Systemen auszutauschen, oder ein System über Kontextinformationen in einem anderen System zu benachrichtigen. Mit dieser Thematik beschäftigen sich nun die folgenden Abschnitte.

4.1 Ort

Ein wichtiger Punkt wenn es um die Modellierung der Kontextinformation Ort geht ist, dass oft eine Echtzeit-Verfolgung von Objekten oder Personen gewünscht wird. Eine weitere grundlegende Anforderung sind die folgenden zwei Abfragen:

- Was ist der Ort eines Objektes?
- Welche Objekte befinden sich ebenfalls an diesem Ort?

Heute existierende Geoinformationssysteme basieren auf statischen Daten in einer Datenbank und sind nicht in der Lage diese Anforderungen zu erfüllen.

Grundsätzlich gibt es zwei verschiedene Modelle um den Ort zu modellieren. Einerseits das symbolische Modell, indem der Ort durch abstrakte Symbole repräsentiert wird. Zum anderen gibt es das geometrische Modell, welches die Ortsinformationen in Form von Koordinaten darstellt.

4.1.2 Symbolisch

Die Wahl des Symbolischen Modells wird getroffen, wenn es sich aufgrund des Ortungssystems anbietet. So benutzen die meisten Systeme, die Ortsinformationen innerhalb von Gebäuden benötigen, das symbolische Modell. Dies hat den einfachen Grund, dass man innerhalb von Gebäuden nicht zufrieden stellend mit geographischen Koordinaten arbeiten kann. So haben fast alle Gebäude mehrere Stockwerke, und selbst wenn es möglich wäre z.B. GPS-Daten zu empfangen, ließe sich nicht feststellen in welchem Raum die Person bzw. das Objekt sich befindet, da zwei übereinander gelegene Räume dieselben GPS-Koordinaten hätten. Die GPS-Höheninformation ist nicht so genau, dass man einzelne Stockwerke von Gebäuden unterscheiden könnte (vorausgesetzt die GPS-Signale könnten innerhalb eines Gebäudes zuverlässig empfangen werden).

4.1.3 Geometrisch

Das geometrische Modell wird verwendet, wenn die Ortungssensoren die Ortsinformationen in Form von Koordinaten liefern. GPS-Koordinaten bestehen beispielsweise aus Längengrad, Breitengrad und Höhe.

4.2 Datenstrukturen

Datenstrukturen werden vom System verwendet um Kontextinformationen zu modellieren, verarbeiten und auszutauschen. Nach [9] fallen diese meistens in eine der folgend vorgestellten Kategorien.

```
<context session="123" action="update">
  <spatial proj="UTM" zone="33" datum="Euro 1950 (mean)">
    <point x="281993" y="4686790" z="205" />>
  </spatial>
</context>
```

Code 4.1

Eine der einfachsten Datenstrukturen sind Schlüssel-Wert-Paare wie man sie beispielsweise von der Klasse *Map* in Java kennt. Schlüssel-Wert-Paare erlauben es, Werte zusammen mit einem Schlüssel abzulegen. Als Schlüssel werden meistens Umgebungsvariablen verwendet, welchen dann ein spezifischer Wert zugeordnet wird. Anhand der Schlüssel ist es möglich die Werte wieder aufzufinden und zu verarbeiten.

Das Projekt ConteXtML [19] verwendet ein einfaches auf XML basierendes Protokoll um Kontextinformationen zwischen einem mobilen Client und einem Server

auszutauschen. Code 4.1 zeigt ein kurzes Beispiel wie Ortsinformationen mit Hilfe von geographischen Koordinaten repräsentiert werden können.

Ein anderes Beispiel wäre das GUIDE-Projekt [15], welches ein objekt-orientiertes Datenstruktur-Modell verwendet. Mit diesem Modell werden Kontextinformationen durch Objektzustände repräsentiert. Ein Objekt wäre beispielsweise ein öffentliches Gebäude in einer Stadt. Der Zustand dieses Gebäudeobjektes besteht aus der aktuellen Belegung verschiedener Variablen. So gibt es z.B. Variablen für die Öffnungszeiten und für Weblinks mit weiterführenden Informationen. Das Objekt bietet Methoden an, um auf diesen Zustand zuzugreifen und ihn zu verändern.

5 Nutzung von Kontext in der Benutzungsschnittstelle

Nachdem in den vorangegangenen Abschnitten Kontext definiert, und untersucht wurde wie man Kontext feststellen und modellieren kann, beschäftigt sich dieser Abschnitt mit der sinnvollen Verwendung von Kontext in Softwaresystemen und Benutzungsschnittstellen. Ziel ist es, durch Verwendung von Kontext für den Benutzer maßgeschneiderte Applikationen zu entwickeln und ihm dadurch einen erhöhten Nutzen zu bieten [8]. Solche Applikationen verändern demnach je nach Kontext ihr Verhalten. So könnte eine Applikation dem Benutzer primär Informationen über seinen aktuellen Aufenthaltsort anbieten und Informationen über weit entfernte Orte erst später präsentieren.

5.1 Anti-Blockier-System

Um den Nutzen und die Nutzung von Kontext in Benutzungsschnittstellen anschaulich zu erörtern, wird im Folgenden das Beispiel des Anti-Blockier-Systems (ABS) beim Automobil untersucht [3].

Das Anti-Blockier-System ist ein kontextbezogenes System welches zur Erhöhung der Sicherheit im Straßenverkehr eingeführt wurde und heute in nahezu allen Automobilen installiert ist. Ziel des Systems ist es, den Bremsweg von Fahrzeugen zu minimieren und die Gefahr des Schleuderns durch falsches Bremsen zu eliminieren. Dazu erhält das System zwei Kontextinformationen als Eingabe, einerseits die Information ob der Fahrer die Bremse betätigt, und andererseits ob die Räder des Automobils aufgrund des Bremsvorgangs blockieren.

Falls dies der Fall ist, verringert das System die Bremskraft solange, bis die Räder sich wieder drehen. Wird in diesem Zustand immer noch die Bremse betätigt erhöht das System wieder die Bremskraft. Dieses System ermöglicht es dem Fahrer durch einfaches Drücken der Bremse schnell und sicher zum Stillstand zu gelangen.

Bevor das System in Automobilen eingebaut war, musste jeder Fahrer die komplexen Zusammenhänge die durch das Zusammenspiel von Bremskraft und Reibung zwischen Reifen und Fahrbahn entstehen, selbst erkennen und beachten. Das System vereinfacht also die mentalen und körperlichen Anforderungen beim Bremsvorgang, es macht die Benutzungsschnittstelle einfacher zu bedienen. Das System übernimmt vom Anwender einen Teil der Kontrolle und vereinfacht das Führen eines Auto-

bils. Des Weiteren vermittelt es dem Fahrer ein vereinfachtes mentales Modell des Bremssystems, bei dem er sich keinerlei Gedanken über zu intensiv angewandte Bremskraft und deren Folgen machen muss.

Neben den Vorteilen die ein solches System zweifelsohne bietet, können anhand des Beispiels aber ebenfalls einige Probleme aufgezeigt werden, die im Zusammenhang mit kontextbezogenen Systemen und Benutzungsschnittstellen entstehen.

Ein potentielles Problem ist, dass das System keinen stabilen Zustand erreicht. Das kann beispielsweise passieren wenn sowohl das System, als auch der Benutzer versuchen sich dem Kontext anzupassen. Im Beispiel des ABS wäre dies der Fall, wenn der Fahrer versucht durch Impulsbremsen ein Blockieren der Räder zu verhindern. Zusätzlich würde jedoch auch das ABS reagieren und der Bremsweg würde sich zwangsläufig verlängern.

Zudem muss immer zwischen Verordnung und Freiheit abgewogen werden. Beim ABS macht es sicherlich Sinn, dass der Fahrer die Räder nicht auf Dauer blockieren kann. Andererseits könnte es doch eine Situation geben in der der Fahrer die Räder blockieren möchte. Dies würde ihm vom ABS versagt.

Schließlich muss der Anwender dem System quasi blind vertrauen. Er muss sich darauf verlassen können, dass das System die nötigen Anpassungen korrekt vornimmt. Dies war bei der Einführung von ABS ein Problem, da viele Fahrer der Technik anfangs nicht trauten.

5.2 Nutzen durch Verwendung von Kontext

Der Hauptnutzen der durch Verwendung von Kontext entsteht ist die vereinfachte Interaktion zwischen dem Anwender und dem System. Diese Vereinfachung betrifft im Wesentlichen drei Punkte:

- Vereinfachung oder Verkleinerung der Aufgabe des Anwenders. Beispielsweise weniger Eingabe durch automatische Erkennung der Absichten des Anwenders.
- Anpassung der Ausgabe des Systems, das heißt Erhöhung der Qualität der Ausgabe bzw. Verringerung der Menge der ausgegebenen Informationen.
- Vereinfachung der Zusammenhänge, die das mentale Modell des Anwenders bestimmen

Schilit beschreibt in [4] vier verschiedene Nutzen, die ein kontextbezogenes System einem Anwender bieten kann. Unter **Auswahlsektion** versteht Schilit eine Technik auf der Benutzungsschnittstelle, welche Objekte die sich in der Nähe befinden hervorhebt oder anderweitig leichter auswählbar macht. Man kann sich beispielsweise eine Auflistung der verfügbaren Drucker vorstellen, welche anhand der Entfernung zum Aufenthaltsort des Anwenders sortiert ist. Der nächste Drucker würde an erster Stelle erscheinen und der am weitesten entfernte an letzter. Da eine solche Benutzungsschnittstelle sehr wahrscheinlich auf einem mobilen Endgerät läuft, müssen die Beschränkungen des Gerätes wie die begrenzte Bildschirmgröße bei der Implementierung beachtet werden. Die beschränkte Größe des Bildschirms erhöht den Nutzen der Auswahlsektion bei mobilen Endgeräten, da dem Anwender die Naviga-

tion in der Benutzungsschnittstelle gewöhnlich schwerer fällt als auf einem stationären Computer mit Tastatur und Maus.

Unter **automatischer kontextbezogener Rekonfiguration** bezeichnet Schilit den Prozess des Hinzufügens neuer Komponenten, Entfernens bestehender Komponenten und das Ändern von Verbindungen zwischen Komponenten. Im obigen Beispiel sind die verfügbaren Drucker die Komponenten.

Kontextbezogene Informationen und Befehle versuchen einem Benutzer in einer bestimmten Situation bestimmte Aktionen zuzuordnen. So führt eine Person in einer Bücherei andere Aktionen aus, als es dieselbe Person in einer Bar tun würde.

Durch Kontext ausgelöste Aktionen sind einfache Wenn-dann-Bedingungen welche festlegen wie die Adaption des Systems erfolgt. Dabei werden diese Aktionen automatisch ausgeführt sobald der entsprechende Kontext auftritt.

	automatisch	manuell
Information	Auswahlsektion & kontextbezogene Information	automatische kontextbezogene Rekonfiguration
Befehl	kontextbezogene Befehle	durch Kontext ausgelöste Aktionen

Tabelle 1: Dimensionen kontextbezogener Software (aus [4])

6 Sicherheit und Privatsphäre

Da Kontext sich auch auf benutzerbezogene und vertrauliche Informationen beziehen kann, muss eine Reglementierung der Datenweitergabe erfolgen. Beispielsweise können Positionsinformationen über eine Person auch missbräuchlich verwendet werden. Andererseits besteht der Nutzengewinn bei vielen kontextbezogenen Systemen eben darin, dass die Kontextinformationen auf einer breiten Basis verfügbar sind.

Durch traditionelle Autorisierungsmethoden ist es möglich nicht-autorisierte Anfragen zu blockieren [10]. Jedoch kann oft erst nach Berechnung eines kontextbezogenen Ergebnisses entschieden werden, wer autorisiert ist das Ergebnis zu erhalten. Ein Beispiel dafür wäre die Anweisung „Veröffentliche nicht die Personen, die in diesem Raum sind.“. Dafür muss ein System allerdings erst einmal berechnen welche Personen in dem Raum sind, wofür es sie lokalisieren muss.

Das Problem ist nicht die eingesetzte Technologie, sondern eher was man mit ihr missbräuchlich machen könnte [9].

Nach [6] achten viele Personen verständlicherweise sehr auf ihre Privatsphäre. Wahrscheinlich ist es niemandem recht, wenn er jederzeit und überall geortet werden kann. Betroffene Personen sollten selbst die Möglichkeit haben zu entscheiden, welche Kontextinformationen sie weitergeben möchten. Menschen wollen Kontrolle darüber haben, was andere Personen über sie erfahren. Sie verhalten sich bezüglich der Weitergabe von Kontextinformationen selektiv.

7 Beispiel-Systeme

Im Folgenden werden einige frühe Beispielsysteme vorgestellt, die Kontext verwenden um einen Zusatznutzen zu erzielen.

7.1 Active Badge Location System

Das „Active Badge Location System“ [16] wurde von der Firma Olivetti in Cambridge (UK) entwickelt. Es war eines der ersten Projekte welches ein kontextbezogenes System entwickelte. Ziel der Applikation ist es, einem/einer Telefonrezeptionist/in bei der Weiterleitung von Anrufen zu unterstützen. Dazu werden sämtliche Angestellte im Haus durch so genannte *Active Badge*“ lokalisiert.

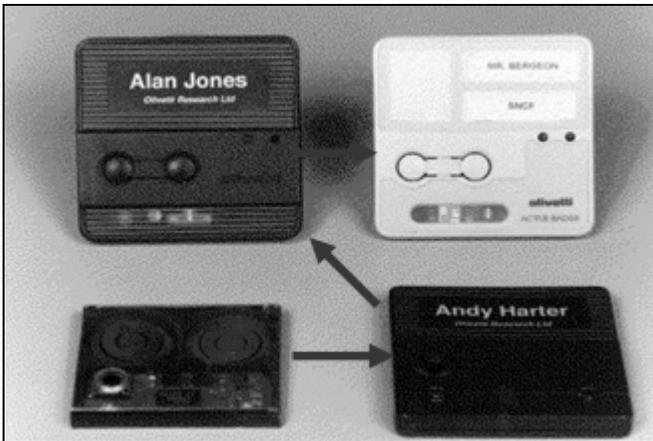


Abbildung 7.1 Active Badge Anhänger (aus [18])

In Abbildung 7.1 sind vier Generationen von *Active Badges* abgebildet. Die *Active Badges* sind Anhänger, den Identifikationsausweisen die in großen Unternehmen getragen werden ähnlich. Die Besonderheit dieser Anhänger ist, dass jedes der Geräte einen Infrarot-Sender enthält, der alle fünfzehn Sekunden ein unterscheidbares Signal aussendet. Diese Signale werden von einer im Haus installierten Infrastruktur von Infrarot-Empfängern aufgenommen. Die Verbindung zwischen Sender und Empfänger muss dabei aufgrund der Infrarot-Technologie nicht unbedingt direkt sein, da die Signale auch von Wänden und anderen Gegenständen reflektiert werden können.

Das auf der Infrastruktur basierende System zeigt einem/einer Telefonrezeptionist/in auf seinem/ihrer Computerbildschirm den Namen eines Angestellten zusammen mit der Telefondurchwahl des nächstgelegenen Telefons und einer kurzen Beschreibung des Ortes an. In Abbildung 7.2 ist die graphische Benutzungsschnittstelle abgebildet. Wie man sieht wird in der dritten Spalte angezeigt mit welcher Wahrscheinlichkeit die Person sich an dem genannten Ort befindet. Ist die Wahrscheinlichkeit kleiner als hundert Prozent, so bewegt sich die Person momentan im Gebäude herum. Ist die Person seit mehr als fünf Minuten nicht mehr lokalisiert wor-

den, so wird die Uhrzeit der letzten Ortung angegeben. Sind mehr als 24 Stunden seit der letzten Ortung vergangen, wird der Tag der letzten Lokalisierung angegeben. Wurde die Person seit über einer Woche nicht mehr geortet so erscheint das Wort „AWAY“ um anzuzeigen, dass die Person sich nicht im Gebäude befindet.

ORL/STL Active Badge Project					
Name	Location	Prob.	Name	Location	Prob.
P Ainsworth	X343 Accs	100%	J Martin	X310 Mc Rm	100%
T Blackie	X222 DVI Rm.	80%	O Mason	X307 Lab	77%
M Chopping	X410 R302	TUE.	D Milway	X307 Drill	AWAY
D Clarke	X316 R321	10:30	B Miners	X202 DVI Rm.	10:40
V Falcao	X218 R435	AWAY	P Mital	X213 PM	11:20
D Garnett	X232 R310	100%	J Porter	X398 Lib.	100%
J Gibbons	X0 Rec.	AWAY	B Robertson	X307 Lab	100%
D Greaves	X304 F3	MON.	C Turner	X307 Lab.	MON.
A Hopper	X434 AH	100%	R Want	X309 Meet. Rm.	77%
A Jackson	X308 AJ	90%	M Wilkes	X300 MW	100%
A Jones	X210 Coffee	100%	I Wilson	X307 Lab.	100%
T King	X309 Meet. Rm.	11:20	S Wray	X204 SW	11:20
D Lioupis	X304 R311	100%	K Zielinski	X402 Coffee	100%

12.00 1st January 1990

Abbildung 7.2 Die graphische Benutzungsschnittstelle des Active Badge Systems (aus [16])

Neben der graphischen Benutzungsschnittstelle können per Kommandozeile die nachfolgend aufgelisteten Aktionen ausgeführt werden:

- **FIND (name):** gibt den momentanen Standort der Person zurück, oder falls sie sich in letzter Zeit bewegt hat, eine Liste der Aufenthaltsorte der vergangenen fünf Minuten zusammen mit der Wahrscheinlichkeit die Person dort anzutreffen
- **WITH (name):** lokalisiert die entsprechende Person und gibt Informationen über Personen in unmittelbarer Umgebung zurück
- **LOOK (location):** gibt eine Liste der Personen zurück, die sich in der Umgebung des angegebenen Ortes befinden
- **NOTIFY (name):** erzeugt einen Alarmton, wenn die Person das nächste Mal lokalisiert wird
- **HISTORY (name):** erzeugt einen kurzen Bericht über die Aufenthaltsorte der Person während der letzten Stunde

Das System ermöglichte es dem/der Telefonrezeptionist/in Anrufe an den, einer Person nächstgelegenen Ort durchzustellen. Das System war so erfolgreich, dass die Telefonrezeptionisten bereits nach einem Monat Schwierigkeiten hatten, wenn das System einmal nicht verfügbar war. Die Angestellten fanden die Weiterleitung von

Anrufen sehr hilfreich, wollten allerdings mehr Kontrolle darüber haben, ob ein Anruf weitergeleitet werden soll oder nicht (z.B. bei einer Besprechung mit dem Vorgesetzten).

7.2 GUIDE

Das GUIDE-System [15] wurde von der Universität Lancaster (UK) entwickelt. Ziel des Projekts ist es ein portables, kontextbezogenes Touristenführungssystem zu entwickeln. Das System ermöglicht es Touristen, beispielsweise aufgrund seiner persönlichen Präferenzen, einen individuell zugeschnittenen Rundgang durch die Stadt Lancaster zu berechnen. Das System gibt dem Touristen Anweisungen, wie er von Punkt A zu Punkt B kommt. Sobald der Tourist an einem für ihn interessanten Ort vorbeikommt erklärt das mobile Endgerät die Sehenswürdigkeit.

Hauptunterschied zu anderen verfügbaren elektronischen Stadtführern ist die drahtlose Kommunikation zwischen mobilem Endgerät und dem System. Dies ermöglicht es den Benutzer durch das mobile Endgerät, Tickets zu buchen oder Rückfragen mit der Touristeninformation zu führen.

Das System besteht im Wesentlichen aus den in Abbildung 7.3 abgebildeten mobilen Endgeräten des Typs *Fujitsu TeamPad 7600* und einer drahtlosen Infrastruktur. Die Endgeräte sind Tablet-PCs mit berührungsempfindlichem Display über das der Benutzer Eingaben vornehmen kann. Die Infrastruktur basiert auf IEEE 802.11 Basisstationen, welche sowohl für die Ortung, als auch für die Datenübertragung verwendet werden.

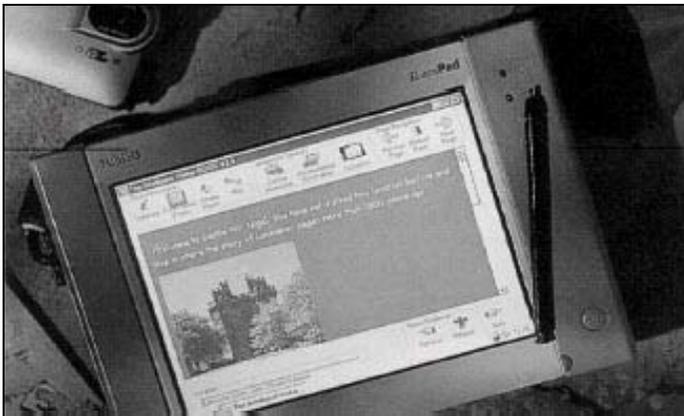


Abbildung 7.3 GUIDE-Endgerät (aus [15])

Beispiele für Kontextinformationen, die das System verwendet sind der aktuelle Ort des Besuchers und der Standort der Sehenswürdigkeiten in der Stadt. Das System reduziert die benötigten Eingaben indem es davon ausgeht, dass die gewünschte Information stark mit dem aktuellen Aufenthaltsort des Besuchers zusammenhängt. So muss beispielsweise ein Tourist, der vor dem Schloss der Stadt steht, nicht umständlich das Schloss aus einer Liste von Sehenswürdigkeiten auswählen. Er wählt einfach den Punkt „*Informationen über die aktuelle Umgebung*“.

Der Vorteil der höheren Qualität der Ausgabe wird beim GUIDE-System beispielsweise erreicht, indem geschlossene oder zuvor bereits besuchte Sehenswürdigkeiten bei einer Auflistung der Sehenswürdigkeiten in der Umgebung, weiter unten in der Liste erscheinen.

Die Vereinfachung des mentalen Modells des Anwenders wird vom System beispielsweise erreicht indem es bei der Berechnung eines Stadtrundgangs eine Vielzahl von Kontextinformationen berücksichtigt. Darunter sind der aktuelle Standort, die Zeit, spezifische Öffnungszeiten von Sehenswürdigkeiten und Präferenzen des Anwenders.

7.3 Context-Call

Context-Call [6] ist eine auf WAP basierende Applikation, die auf Mobiltelefonen ausgeführt wird. Ziel der Applikation ist es dem Benutzer zu ermöglichen potentiellen Anrufern seinen Kontext mitzuteilen. Basierend auf dieser Kontextinformation kann der Anrufer nun entscheiden, ob es gerade passend ist die Person anzurufen.

Die Applikation möchte erreichen, dass der Mobilfunknutzer weder wichtige Anrufe verpasst (nämlich wenn er das Mobiltelefon ausschaltet), noch dass er durch unpassende Anrufe belästigt wird (z.B. in einer Besprechung).

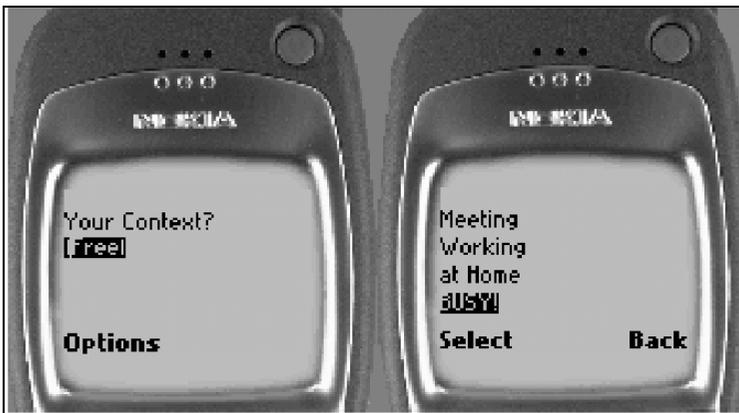


Abbildung 7.4 Context-Call (aus [6])

In Abbildung 7.4 ist die Benutzungsschnittstelle von Context-Call zu sehen, es besteht die Möglichkeit seinen derzeitigen Kontext einzugeben. Versucht nun eine andere Person via Context-Call eine Verbindung zum Teilnehmer aufzubauen, so wird ihm zuerst der Kontext des Angerufenen mitgeteilt. Daraufhin gibt es die drei Auswahlmöglichkeiten „Anrufen“, „Nachricht auf Mailbox hinterlassen“ und „Anruf abbrechen“.

Im normalen Leben werden Personen die gerade sehr beschäftigt sind nur angesprochen, wenn es sehr wichtig ist. Mensch-Zu-Mensch-Kommunikation in einem sozialen Umfeld ist reichhaltiger als Kommunikation über das Telefon oder den Computer. Grund dafür ist der Kontext, den man bei einem Gespräch gegenseitig teilt, beispielsweise Grimassen die über den Gefühlszustand Auskunft geben, oder einfach

der Aufenthaltsort. Context-Call versucht die Telekommunikation der normalen Mensch-Zu-Mensch-Kommunikation anzunähern.

7.4 Orientierungs-sensibler PDA

Bei diesem Projekt [1] wurde ein *Apple Newton MessagePad* PDA mit einem Orientierungssensor ausgestattet. Dieser Sensor besteht aus zwei Quecksilberschaltern. Das Betriebssystem des PDA unterstützt eine Funktion, die es ermöglicht das Display in alle vier Richtungen zu schalten, d.h. der Benutzer kann den Bildschirm sowohl vertikal, als auch horizontal ablesen. Diese Funktion wurde nur von sehr wenigen Anwendern benutzt. Ziel des Projektes war es nun, anhand der ermittelten Orientierung (dem Kontext), das Display automatisch richtig auszurichten.

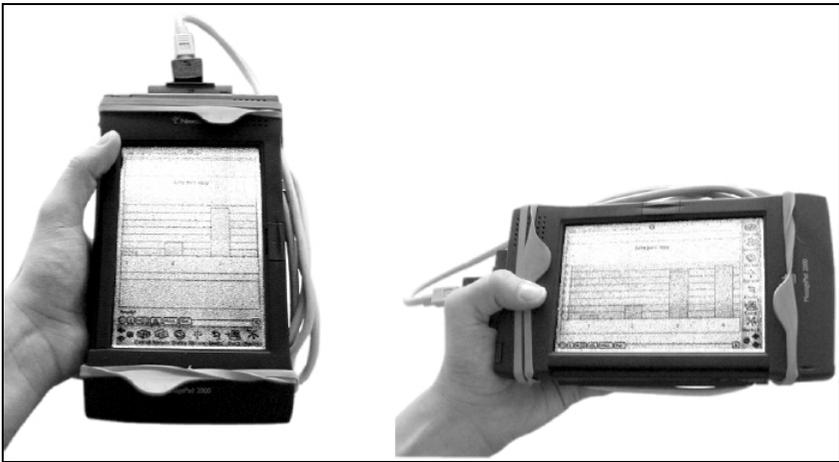


Abbildung 7.5 Orientierungs-sensibler PDA (aus [1])

Abbildung 7.5 zeigt zwei unterschiedliche Ausrichtungen des Bildschirms, je nach Orientierung des PDA. Bei diesem System besteht die reduzierte Eingabe des Benutzers darin, dass er die Ausrichtung des Bildschirms nicht mehr per Hand umstellen muss. Das erhöht den Komfort und lässt den Benutzer die Funktion voraussichtlich öfter verwenden.

Eine kommerzielle Abwandlung dieses Systems findet man in modernen Digitalkameras. Einige Modelle richten die Bilddateien der geschossenen Aufnahmen automatisch richtig aus, indem sie die Orientierung der Kamera zum Auslösezeitpunkt als Kontext verwenden.

8 Zusammenfassung

Die große Anzahl verschiedener Definitionen des Begriffs Kontext zeigt, dass es sich um ein noch sehr junges Forschungsgebiet handelt. In Zukunft werden sich die

Forscher wohl auf eine konkrete Definition einigen müssen, um die Entwicklung kontextbezogener Systeme zu erleichtern.

Die Modellierung von Kontext geschieht bis heute noch sehr uneinheitlich. Grundsätzlich hat bis heute jedes kontextbezogene System seine eigene Art der Modellierung implementiert. Doch um kontextbezogene Systeme leicht implementierbar zu machen und einen Austausch von Kontextinformationen zwischen verschiedenen Systemen zu ermöglichen, kommt man wohl kaum um die Entwicklung von allgemein akzeptierten Standards herum.

Anschließend haben wir gesehen, dass es bereits heute einen erhöhten Nutzen durch Verwendung von kontextbezogenen Benutzungsschnittstellen gibt. Dies weckt hohe Erwartungen an die Zukunft. Die Technologien und Systeme werden in der Zukunft immer komplexer, wodurch sich auch die Schwierigkeit erhöht, solche Systeme zu bedienen und deren Möglichkeiten auszuschöpfen. Kontextbezogene Benutzungsschnittstellen sind ein wichtiges Instrument, um anwendungsfreundliche Systeme zu implementieren. Fortschritte auf dem Gebiet der kontextbezogenen Benutzungsschnittstellen werden Systeme in Zukunft leistungsfähiger und leichter bedienbar machen.

Das Thema Sicherheit und Privatsphäre sollte hierbei allerdings nicht unterschätzt werden. Obwohl es bisher wenig Forschung zu diesem Thema gibt, spielt es eine essentielle Rolle für den Erfolg von kontextbezogenen Systemen. Denn nur wenn die betroffenen Personen, den Systemen und der Diskretion bei der Weitergabe personenbezogener Kontextinformationen vertrauen können, steht einer breiten Akzeptanz solcher Systeme nichts mehr im Wege.

Literatur

- [1]: Schmidt, A., Beigl, M., Gellerson, H.-W.. There is more to Context than Location - Environment Sensing Technologies for Adaptive Mobile User Interfaces. Proceedings of Workshop on Interactive Applications of Mobile Computing. 1998
- [2]: Paymans, T., Lindenberg, J., Neerincx, M.. Usability Trade-offs for Adaptive User Interfaces: Ease of Use and Learnability. Proceedings of the 9th international conference on Intelligent user interface. Portugal. 2004
- [3]: Cheverst, K., Davies, N., Mitchell, K., Efstratiou, C.. Using Context as a Crystal Ball: Rewards and Pitfalls. Proceedings of Workshop on Situated Interaction in Ubiquitous Computing. 2000
- [4]: Schilit, B., Adams, N., Want, R.. Context-Aware Computing Applications. IEEE Workshop on Mobile Computing Systems and Applications. 1994
- [5]: Dey, A.. Understanding and Using Context. Personal and Ubiquitous Computing. Volume 5, Issue 1. Pages: 4 – 7. 2001
- [6]: Schmidt, A., Takaluoma, A., Mäntyjärvi, J.. Context-Aware Telephony over WAP. Handheld and Ubiquitous Computing (HUC2k). 2000
- [7]: Brown, P., Bovey, J., Chen, X.. Context-aware Applications: from the Laboratory to the Marketplace. IEEE Personal Communications, 4(5):58–64, 1997
- [8]: Dey, A., Abowd, G.. The Context Toolkit: Aiding the Development of Context-Aware Applications. Seiten 434-441. 1999
- [9]: Chen, G., Kotz, D.. A Survey of Context-Aware Mobile Computing Research. Technical Report: TR2000-381. 2000

- [10]: Leonhardt, U., Magee, J., Dias, P.. Location Service in Mobile Computing Environments. In International Workshop on Visualization and Mobile Computing, 1996.
- [11]: Dey, A., Abowd, G.. Towards a Better Understanding of Context and Context-Awareness. In Workshop on The What, Who, Where, When, and How of Context-Awareness, Conference on Human Factors in Computer Systems. 2000
- [12]: Schilit, B., Theimer, M. Disseminating Active Map Information to Mobile Hosts. IEEE Network, 8(5) 22-32. 1994
- [13]: Brown, P.J., Bovey, J.D. Chen, X. Context-Aware Applications: From the Laboratory to the Marketplace. IEEE Personal Communications, 4(5) 58-64. 1997
- [14]: Ryan, N., Pascoe, J., Morse, D. Enhanced Reality Fieldwork: the Context-Aware Archaeological Assistant. Gaffney, V., van Leusen, M., Exxon, S. (eds.) Computer Applications in Archaeology. 1997
- [15]: K. Cheverst, N. Davies, K. Mitchell, A. Friday and C. Efstratiou. Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. Proceedings of CHI 2000, Netherlands, pp 17-24. 2000
- [16]: R. Want, A. Hopper, V. Falcão, J. Gibbons. The Active Badge Location System. Olivetti Research Ltd. (ORL), Cambridge, England. ACM Transactions on Information Systems (TOIS). Volume 10, Issue 1. 1992
- [17]: Weiser, M., The computer for the 21st century. Scientific American, 94-104, September 1991
- [18]: AT&T Laboratories, Cambridge. The Active Badge. 02.02.2005.
<<http://www.uk.research.att.com/thebadge.html>>
- [19]: Nick Ryan. ConteXtML: Exchanging Contextual Information between a Mobile Client and the FieldNote Server. 02.02.2005.
<<http://www.cs.kent.ac.uk/projects/mobicomp/fnc/ConteXtML.html>>

Autorenverzeichnis

Bahr, Julius	15
Block, Florian	55
Breisinger, Marc	137
Ecker, Ronald	1
Falke, Oliver	239
Freund, Stefan.....	189
Kern, Dagmar	71
Otto, Friederike.....	207
Ruseva, Radostina.....	173
Schicker, Matthias	35
Schlerf, Benjamin	157
Schrittenloher, Martin.....	105
Wimmer, Raphael.....	117
Wnuk, Monika	87
Zelhofer, Nora.....	223

